Optimal liquidation in dark pools in discrete and continuous time

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium (Dr. rer. nat.)
im Fach Mathematik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät II
Humboldt-Universität zu Berlin

von
Dipl.-Math. Peter Kratz
geboren am 8. Juli 1980 in Ulm

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr. Jan-Hendrik Olbertz

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät II:
Prof. Dr. Elmar Kulke

Gutachter:
1. Prof. Dr. Ulrich Horst
2. Prof. Dr. Stefan Ankirchner
3. Prof. Dr. Darrell Duffie
4. Prof. Dr. Ivar Ekeland

eingereicht am: 12. Februar 2011
Tag der Verteidigung: 12. Juli 2011

Abstract

In recent years there has been an increasing interest in financial market models that account for the impact of large transactions on asset prices. In this thesis we consider an "illiquid market" with a risk-averse investor who has to liquidate a large portfolio within a finite time horizon $[0, T]$. At any point in time, the investor has the option to trade at a traditional exchange (the "primary venue") which yields price impact and to place orders in a so-called *dark pool*.

The liquidity in dark pools is not openly displayed and dark pools do not contribute to the price formation process. Instead, orders are executed at the price of the primary venue if matching liquidity is available. Therefore, these orders have no price impact. However, in contrast to the primary venue, their execution is uncertain. The investor thus faces the trade-off between the direct costs resulting from the price impact of her orders at the primary venue and the indirect costs resulting from the execution uncertainty in the dark pool. This thesis considers stochastic control problems arising in portfolio liquidation problems using dark pools.

At first we consider a discrete-time market model and a cost functional which incorporates both the expected price impact costs and the market risk of a large portfolio. For linear temporary price impact, this results in a linear-quadratic cost functional and we obtain the optimal trading strategy in the form of an explicit backward recursion. If a single asset position is to be liquidated, the investor trades a certain proportion of the position at the primary venue, with the remainder being placed in the dark pool. For multi asset liquidation this is generally not optimal, and the optimal strategy depends strongly on the correlation of the assets. We generalize this model for single asset liquidation and analyze *adverse selection* effects. This reduces the attractiveness of the dark pool and thus optimal dark pool orders are smaller; it can even be optimal to not use the dark pool at all.

In the second part of the thesis we consider the counterparts of the described optimization problems in a continuous-time market model. In continuous time the liquidation constraint implies a singularity of the value function at the terminal time T. Via the HJB equation and a quadratic ansatz, we obtain a candidate for the value function of the linear-quadratic optimization problem without adverse selection. This candidate is the limit of a sequence of solutions of initial value problems for a matrix differential equation. Although the differential equation is not a Riccati equation, we are able to show that this limit actually exists by using an appropriate matrix inequality and a known comparison result for Riccati equations. Additionally, we thereby obtain upper and lower bounds of the solutions of the initial value problems, which enables us to prove a verification theorem. The complexity of the optimization problem with adverse selection stems form the non-linear-quadratic form of the cost functional. Via the HJB equation and extensive heuristic considerations we prove that the value function is a "quasi-polynomial" whose coefficients depend on the asset position in a non-trivial way.

For single asset liquidation, all optimization problems are solved in closed form both in discrete and in continuous time. The solutions of the discrete-time optimization problems converge to the solutions of the continuous-time optimization problems under appropriate conditions.

Zusammenfassung

In mathematischen Finanzmarktmodellen wird seit einigen Jahren zunehmend der Einfluss großer Transaktionen auf die Preisbildung von Wertpapieren berücksichtigt. In dieser Arbeit soll nun ein solcher „illiquider Markt“ betrachtet werden, auf dem ein risikoaverser Investor ein großes Portfolio bis zu einem festen Zeitpunkt T liquidieren muss. Der Investor hat dabei zu jedem Zeitpunkt die Möglichkeit, gleichzeitig auf einem traditionellen Markt (dem „Primärmarkt“) zu handeln, wobei er den Preis beeinflusst, und Aufträge in einem sogenannten *Dark Pool* zu erteilen.

Die Liquidität in Dark Pools ist nicht öffentlich bekannt, und diese Märkte nehmen nicht an der Preisfindung teil. Vielmehr werden Aufträge zum Preis des Primärmarkts abgewickelt, sofern eine Transaktion möglich ist. Deshalb haben diese Aufträge zwar keinen Preiseinfluss, deren Ausübung ist aber im Gegensatz zum Primärmarkt nicht garantiert. Der Händler muss also zwischen den direkten Kosten, die durch den Markteinfluss seines Handelns auf dem Primärmarkt und den indirekten Kosten, die durch die Ausübungsunsicherheit seiner Aufträge im Dark Pool entstehen, abwägen. Ziel der Arbeit ist die Lösung von stochastischen Steuerungsproblemen, die im Rahmen eines Modells auftreten, das diesen Zielkonflikt möglichst realistisch darstellt.

Zunächst betrachten wir in einem zeitdiskreten Handelsmodell ein Kostenfunktional, das sowohl die erwarteten Kosten durch den Preiseinfluss als auch das Marktrisiko einer großen Position berücksichtigt. Für linearen temporären Preiseinfluss erhalten wir ein linear-quadratisches Kostenfunktional und können die optimale Handelsstrategie in Form einer expliziten Rekursion bestimmen. Liquidiert der Investor eine Position in einem einzelnen Wertpapier, so baut er sie langsam auf dem Primärmarkt ab und bietet gleichzeitig den Rest im Dark Pool an. Besteht die Position aus mehreren Wertpapieren, ist dies im Allgemeinen nicht mehr optimal, und die optimale Strategie hängt stark von der Korrelation der Wertpapiere ab. Für die Liquidation in einem einzelnen Wertpapier verallgemeinern wir dieses Modell und untersuchen die Auswirkungen von *Adverser Selektion*. Diese verringert die Attraktivität des Dark Pools und führt dazu, dass ein kleinerer Teil der Position im Dark Pool angeboten wird oder dieser gar nicht genutzt wird.

Im zweiten Teil der Arbeit betrachten wir ein zeitstetiges Handelsmodell und in diesem die Entsprechungen der beiden beschriebenen Optimierungsprobleme. In stetiger Zeit impliziert die Liquidationsbedingung eine Singularität der Wertfunktion am Endzeitpunkt T. Mittels der HJB Gleichung und eines quadratischen Ansatzes erhalten wir einen Kandidaten für die Wertfunktion des linear-quadratischen Liquidationsproblems ohne Adverse Selektion. Dieser wird beschrieben durch den Grenzwert einer Folge von Lösungen einer Matrix Differentialgleichung. Obwohl diese Differentialgleichung keine Riccati Gleichung ist, können wir mit Hilfe einer geeigneten Matrixungleichung und eines bekannten Vergleichsarguments für Riccati Gleichungen beweisen, dass dieser Grenzwert existiert. Außerdem erhalten wir dadurch obere und untere Schranken der Lösungen der Anfangswertprobleme, die es uns ermöglichen, ein Verifikationsargument durchzuführen. Die Schwierigkeit des Optimierungsproblems mit Adverser Selektion liegt in der Struktur des Kostenfunktionals. Dieses ist nicht linear-quadratisch. Mittels der HJB Gleichung und umfangreichen heuristischen Betrachtungen gelingt es uns zu beweisen, dass die Wertfunktion ein quadratisches „Quasi-Polynom“ ist, dessen Koeffizienten in nicht-trivialer Weise von der Position abhängen.

Liquidiert der Investor eine Position in einem einzelnen Wertpapier, so können wir sowohl im zeitdiskreten als auch im zeitstetigen Fall sämtliche Optimierungsprobleme in geschlossener Form lösen. Unter geeigneten Bedingungen konvergieren die Lösungen der diskreten Optimierungsprobleme gegen die Lösungen der stetigen Optimierungsprobleme.

Contents

Introduction

In the last years equity trading has been transformed by the advent of so-called *dark pools*. These alternative trading venues differ significantly from classical exchanges. Especially in the US equities market dark pools have gained a significant market share increasing to about 8%. Although dark pools vary in a number of properties such as ownership, crossing procedure and accessibility (see Mittal [2008] for further details and a typology of dark pools), they generally share the following two stylized facts. First, dark pools are dark. The liquidity available is not quoted, making trade execution uncertain and unpredictable. Second, dark pools do not determine prices. Instead, they monitor the prices determined by the classical exchanges and settle trades in the dark pool only if possible at these prices. Thus, trades in the dark pool have no or less price impact.[1] This thesis is concerned with the solution of stochastic optimal control problems in discrete and continuous time arising in the context of optimal portfolio liquidation if a large investor has access both to a classical exchange (or "primary venue" or "primary exchange") and to a dark pool.[2] More precisely, we study a model for optimal liquidation of a large portfolio consisting of n assets within a finite time horizon $[0, T]$ reflecting the trade-off between execution uncertainty of dark pool orders and price impact costs of trading at the primary venue. To our best knowledge, this is the first mathematical framework within which optimal trade execution for a large multi asset portfolio using a classical exchange and a dark pool simultaneously is analyzed.

In Chapter 1 we introduce a discrete-time setting where trading is possible at finitely many points in time $0 = t_1 < \cdots < t_N = T$. In a first step we provide existence and uniqueness of optimal liquidation strategies for a general form of price impact and dark pool liquidity by a non-standard convexity argument. In a second step we consider linear price impact and assume that dark pool orders are executed fully or not at all in order to obtain further insights into the structure of the optimal strategy and the value function of the optimization problem. If future price moves are independent of current dark pool liquidity, the resulting cost functional is of linear-quadratic form and we derive solutions of the quadratic value function and the linear optimal strategy in the form of explicit backward recursions for multi asset liquidation ($n \in \mathbb{N}$) and in closed form for single asset liquidation ($n = 1$). For $n = 1$, a certain proportion of the position is liquidated at the exchange while the remainder is placed in the dark pool. This is

[1] For empirical evidence for lower transaction costs or price impact of dark pools compared to classical exchanges see, e.g., Conrad et al. [2003] and Fong et al. [2004].

[2] The overall liquidity traded in dark pools in the US is strongly fragmented among approximately 40 different venues, see e.g., Carrie [2008]. Therefore, liquidity aggregation is a major issue. Ganchev et al. [2010] and Laruelle et al. [2009] establish learning algorithms to achieve optimal order split *between* dark pools. Instead of analyzing the simultaneous use of several dark pools, we consider such an "aggregated" dark pool in our model which we call "the dark pool".

no longer optimal for $n \geq 2$. The optimal orders both in the primary exchange and in the dark pool depend strongly on the correlation of the assets. In particular, it can be optimal to place orders in the dark pool that are considerably smaller or larger than the remainder of the position.

Subsequently, we generalize the linear price impact setting for the case $n = 1$ in order to allow for "adverse selection" effects. Under adverse selection, the value function is not of linear-quadratic form and the optimal strategy is not linear. We derive the structure of the value function heuristically and solve the optimization problem in closed form. Adverse selection reduces the attractiveness of the dark pool, and it is no longer optimal to place the full remainder of the position in the dark pool. It can even be optimal to not use the dark pool at all. The discrete-time setting of Chapter 1 is based on joint work with Torsten Schöneborn, see Kratz and Schöneborn [2010].

In Chapters 2 and 3 we consider optimal liquidation with linear price impact *with* respectively *without* adverse selection in continuous time. This gives rise to continuous-time stochastic optimal control problems with jumps. Due to the liquidation constraint, the value functions of the problems have singularities at the terminal time T, and hence the verification arguments require non-standard considerations.

In either case, the dark pool liquidity is modeled by an n-dimensional Poisson process. In the absence of adverse selection, the cost functional is linear-quadratic. We approximate the liquidation constraint by a sequence of modified optimization problems with increasing finite end-costs. Via a quadratic ansatz, the corresponding Hamilton-Jacobi-Bellman (HJB) equation suggests that the value functions of the modified optimization problems are quadratic forms for matrix-valued functions which are the solutions of initial value problems for a specific matrix differential equation. Explicit solutions for these initial value problems are not known, and the differential equation is not a Riccati matrix differential equation. We establish a matrix inequality that allows us to apply a known comparison result for Riccati matrix differential equation in order to obtain closed form upper and lower bounds for the solutions of the initial value problems. The bounds are important for several reasons. First, they enable us to prove a verification theorem for the modified optimization problems with finite end-costs. Second, they imply the existence of the limit of the solution of these optimization problems and thus yield a well-defined candidate for the value function of the original optimization problem with liquidation constraint. Third, the limit of the bounds transfers to bounds for this candidate value function and we deduce a verification theorem for the solution of the optimal liquidation problem.

In Chapter 3 we analyze adverse selection in a *single asset* model. The resulting optimization problem is not of linear-quadratic form and the value function is not quadratic. Via extensive heuristic considerations, we obtain candidates for the value function and the optimal strategy in closed form. These are a quadratic "quasi-polynomial" respectively a "quasi-affine linear" function with the coefficients depending on the state variable in a non-trivial way. The candidate for the value function is well-defined and differentiable and we are thus able to prove a verification theorem.

The thesis contributes to two lines of research in economic theory and mathematical finance. A first line takes the liquidity effects as given and derives optimal trading strategies within a stylized market model. Several such models have recently been proposed for classical exchanges, e.g., Bertsimas and Lo [1998], Almgren and Chriss [2001], Almgren [2003], Obizhaeva and Wang [2006], Schied and Schöneborn [2009], Alfonsi et al. [2010] and Schied et al. [2010]. None of these models includes the effects of dark pools.

A second line of research focuses on the underlying mechanisms for illiquidity effects and models the competition between financial markets when assets are traded at different markets and one of the markets is a dark pool. Hendershott and Mendelson [2000] analyze the interaction of dealer markets and a crossing network in a static one period framework, where each investor buys or sells a *single* share. Degryse et al. [2009] introduce a dynamic multi-period framework. Our research differs from these models for several reasons. None of these models captures the effect of price impact of a single large trader and the possibility of dynamically splitting orders between the primary market and the dark pool. By contrast, we take the position of a single trader and consider price moves, dark pool liquidity and price impact as given exogenously. Thus, we are able to incorporate price impact in a general form and allow for adverse selection. Furthermore, we consider a multi asset dynamic setting where the correlation between stocks matters for risk-averse traders.

In continuous time the liquidation constraint yields a singularity of the value function at the terminal time T, and thus the resulting stochastic control problems require non-standard considerations. For optimal liquidation without dark pools for CARA investors, such an optimization problem with singularity at the terminal time T has been studied by Schied et al. [2010]. In this case the optimal control problem does not include jumps.

For optimal liquidation *without* adverse selection, the optimization problem is linear-quadratic. For $n = 1$ and finite end-costs (hence no singularity of the value function at time T), the solution of the control problem is well-known, see, e.g., Øksendal and Hu [2008]. Naujokat and Westray [2010] and Höschler [2011] have solved the *single asset* case with infinite end-costs contemporaneously with our work, however in different contexts. Naujokat and Westray [2010] use the stochastic maximum principle rather than HJB equations for the solution. The difficulty in our setting stems form the *combination* of the singularity of the value function and the fact that we consider multi-dimensional portfolios; thus the solution of the optimization problem involves the detailed analysis of a specific non-Riccati-type matrix differential equation, for which we establish existence results and upper and lower bounds of the solution by means of a novel matrix inequality.

For optimal liquidation *with* adverse selection the difficulty originates form the non-linear-quadratic form of the cost functional. The main contribution is therefore the fact that we provide the solution in closed form.

Summary of the discrete-time part

In Chapter 1 we introduce a discrete-time market model for a large investor who aims to liquidate a given portfolio $X(t_0) \in \mathbb{R}^n$ of n possibly correlated assets in finitely many time steps $0 = t_0 < \cdots < t_N = T$. We consider price moves, dark pool liquidity and price impact as given exogenously. The transaction price $P(t_i) \in \mathbb{R}^n$ at the primary exchange at time t_i can be decomposed into the price impact of the primary venue trades $(x_j)_{j=0,\dots,i}$ of the large investor and the "fundamental" asset price $\tilde{P}(t_i)$ that would have occurred in the absence of large trades:

$$P(t_i) = \underset{\substack{\uparrow \\ \text{“Fundamental”} \\ \text{asset price}}}{\tilde{P}(t_i)} \quad - \quad \underset{\substack{\uparrow \\ \text{Price impact} \\ \text{of investor}}}{f_i(x(t_0), \dots, x(t_i))}.$$

We assume that $\tilde{P}$ is a stochastic process with independent increments $\epsilon(t_i) \in \mathbb{R}^n$ and that the overall price impact costs of trading

$$\sum_i x(t_i)^\top f_i(x(t_0), \dots, x(t_i))$$

are strictly convex and grow superlinearly in the strategy (Assumption 1.1.1). Our price impact model extends, e.g., the models of Almgren and Chriss [2001] and Obizhaeva and Wang [2006].

The liquidity in the dark pool at time t_i is given by random variables $a(t_i)$ and $b(t_i) \in \mathbb{R}^n$ (for the supply respectively the demand) independent of previous dark pool liquidity and price moves (Assumption 1.1.3). Trades $y(t_i)$ in the dark pool are executed until the liquidity is exhausted and the executed part of the order $y(t_i)$ is denoted by $z(t_i)$:

$$z_k(t_i) := \begin{cases} \min(y_k(t_i), b_k(t_i)) & \text{if } y_k(t_i) \geq 0 \\ -\min(-y_k(t_i), a_k(t_i)) & \text{if } y_k(t_i) < 0, \end{cases}$$

where the subscript k is used for asset k, and positive and negative orders denote sell respectively buy orders. Notice that in this general setting the dependence of $z(t_i)$ on the order $y(t_i)$ is non-linear. While the investor's trades in the primary exchange have a feedback effect on the market price, the price in the dark pool is the unaffected fundamental asset price $\tilde{P}$. If we replace the transaction price in the dark pool by P, market manipulation can become profitable and optimal trading strategies might not exist (cf. Section 1.5).

At any point in time t_i, the risk-averse investor can choose among all adapted strategies $(x, y) = (x(t_i), \dots, x(t_N), y(t_i), \dots, y(t_N))$ that liquidate the given portfolio $X(t_i)$ (Definition 1.1.4):

$$\sum_{j=i}^{N} x(t_j) + z(t_j) = X(t_i).$$

Her trading objective is to minimize the following cost functional:

$$\underbrace{\mathbb{E}\Big[X(t_i)^\top \tilde{P}(t_i) - \sum_{j=i}^{N} \big(x(t_j)^\top P(t_j) + z(t_j)^\top \tilde{P}(t_j)\big)\Big]}_{\text{Expected implementation shortfall}} + \underbrace{\alpha \mathbb{E}\Big[\sum_{j=i}^{N} X(t_j)^\top \Sigma(t_{j+1}) X(t_j)\Big]}_{\text{Risk costs}}.$$

Here, $X(t_j)$ is the portfolio position at time t_j associated with the strategy (x, y), the matrix $\Sigma(t_{j+1})$ denotes the covariance matrix of the price increments of the n assets and $\alpha \geq 0$ is the risk-aversion parameter of the investor. The criterion penalizes market risk and hence slow liquidation. In the setting of Almgren and Chriss [2001] for optimal liquidation without dark pools this is equivalent to minimizing a mean-variance functional over all deterministic strategies, which is in turn equivalent to minimizing the utility of a CARA investor over all strategies by Schied et al. [2010].

In Section 1.2 we consider the general form of price impact and dark pool liquidity above and assume additionally that dark pool liquidity is independent across the different assets (Assumption 1.2.1). Despite the general form of the model, we are able to prove that there exists a unique optimal liquidation strategy (Theorem 1.2.2). Due to the non-linear dependence of the executed fraction $z(t_i)$ on the dark pool order $y(t_i)$, proving uniqueness via a convexity argument requires non-standard considerations (Theorem 1.2.4).

In Section 1.3 we consider a more specific setting in order to obtain further insights into the structure of the value function of the optimization problem and the optimal liquidation strategy. We assume that the price impact of the investor is linear and temporary given by a positive definite matrix Λ and that orders in the dark pool are executed fully or not at all (Assumption 1.3.1 (i) and (iii)). This market impact model goes back to the work of Almgren and Chriss [2001] that has become the basis of several theoretical studies, e.g., Rogers and Singh [2010], Almgren and Lorenz [2007], Carlin et al. [2007] and Schöneborn and Schied [2009]. Additionally, it demonstrated reasonable properties in real world applications and serves as the basis of many optimal execution algorithms run by practitioners (see, e.g., Kissell and Glantz [2003], Schack [2004], Abramowitz [2006] and Leinweber [2007]). We also assume that future price moves in the market are independent of the current liquidity in the dark pool and that the fundamental asset price is a martingale (Assumption 1.3.1 (ii) and (iii)). The optimization problem then simplifies to minimizing the linear-quadratic functional

$$\mathbb{E}\Big[\sum_{j=i}^{N} x(t_j)^\top \Lambda x(t_j)\Big] + \alpha \mathbb{E}\Big[X(t_j)^\top \Sigma X(t_j)\Big]. \tag{1}$$

By backward induction, we derive recursions for the value function and the optimal strategy and show that these are quadratic respectively linear in the portfolio position (Theorem 1.3.4). Thus, we are able to analyze the properties of the optimal strategy in detail. In a single asset setting, a certain proportion of the position is traded at the primary exchange while at the same time the remainder of the position is being offered in the dark pool. Compared to optimal liquidation without dark pools, the trading at the primary exchange is "slower". We solve this case in closed form and deduce monotonicity properties for the optimal strategy and the components of the costs of the optimal strategy. If a multi asset portfolio is to be liquidated, the optimal liquidation strategy depends strongly on the correlation of the n assets, and it is no longer optimal to offer the whole remainder of the portfolio in the dark pool. If, e.g., the portfolio is balanced and thus only exposed to little market risk, then a complete liquidation of the position in one of the assets is unfavorable and thus only a fraction of the entire portfolio should be placed in the dark pool. This emphasizes that overly simple adjustments of existing trading algorithms in order to include dark pools can yield undesirable effects.

While there is no feedback from the dark pool to the price determined at the exchange, the two venues can be connected since liquidity in the dark pool and price movements at the primary exchange can be dependent. For example, liquidity on the bid side of the dark pool might be unusually high exactly when prices move up. We call the phenomenon that a trader's order is executed in the dark pool shortly before an impending favorable price move in the market "adverse selection".

Adverse selection in dark pools has been a topic both in theoretical and in empirical research. In their equilibrium model, Hendershott and Mendelson [2000] find that adverse selection occurs endogenously due to information asymmetry in the market. Traders with less information about future price moves risk to be adversely selected in the dark pool by well-informed traders. Næs and Ødegaard [2006] analyze transaction costs of crossing networks empirically and find that potential savings are possibly mitigated by adverse selection. Using the same data set, Næs and Skjeltorp [2003] find alternative explanations for these effects and confirm that in spite of possible adverse selection, transaction costs in crossing networks are lower than in the primary market.

Our general market model allows for the dependence of future price moves on dark pool liquidity, and we can thus incorporate and analyze adverse selection in Section 1.4. To this end, we generalize the linear price impact setting of Section 1.3 for the case of single asset liquidation and model adverse selection by assuming that

$$\mathbb{E}[\epsilon(t_{i+1})|a(t_i)=\infty] =: -\Gamma < 0, \quad \mathbb{E}[\epsilon(t_{i+1})|b(t_i)=\infty] =: \Gamma > 0$$

(Assumption 1.4.4). The resulting cost functional is then given by

$$\mathbb{E}\Big[\sum_{j=i}^{N} \left(\Lambda x(t_j)^2 + p\Gamma|y(t_j)|\right)\Big] + \alpha\mathbb{E}\Big[\Sigma X(t_j)^2\Big], \tag{2}$$

where p is the probability of dark pool execution. This can be interpreted as "linear" costs for dark pool orders. Because of the terms $p\Gamma|y(t_j)|$, this is not a linear-quadratic functional, and we cannot expect the value function to be quadratic. We derive the structure of the value function from extensive heuristic considerations and obtain explicit recursions for the optimal strategy and the value function. The value function is piecewise a quadratic polynomial, and the optimal strategy is piecewise affine linear in the asset position (Theorem 1.4.4). We also establish closed form solutions for the value function and the optimal strategy (Corollary 1.4.10) and analyze their properties (Section 1.4.5). Adverse selection lowers the attractiveness of the dark pool relative to the primary exchange and changes the structure of the optimal strategy and the value function significantly. It is no longer optimal to place the entire remainder of the position in the dark pool. For small positions, it can even be optimal to not use the dark pool at all.

Summary of the continuous-time part

In Chapters 2 and 3 we transfer the discrete-time trading model with linear price impact of Chapter 1 into a continuous-time trading model. Instead of placing orders $x(t_i)$ ($i = 0, \dots, N$) in the traditional exchange, the investor can control her trading intensity $\xi(s)$ ($s \in [0,T)$). We model trade execution of orders $\eta(s)$ in the dark pool by n independent Poisson processes $\pi_1, \dots, \pi_n$ with intensities $\theta_1, \dots, \theta_n \geq 0$.

Given a point in time $t \in [0,T)$, a portfolio position $x \in \mathbb{R}^n$ at time t and a trading strategy $u = (\xi, \eta)$, the corresponding portfolio position $X^u(s)$ at time $s \in [t,T)$ is hence given by the solution of the following stochastic differential equation:

$$\begin{aligned} dX^u(s) &= -\xi(s)ds - \eta(s)d\pi(s) \\ X^u(t) &= x, \end{aligned}$$

where $\pi = (\pi_1, \dots, \pi_n)^\top$.

All admissible liquidation strategies must satisfy the *liquidation constraint*

$$\lim_{s \to T-} X^u(s) = 0 \tag{3}$$

(Definition 2.1.1). Therefore, the value functions v of the optimization problems we consider have singularities at the terminal time T:

$$\lim_{s \to T-} v(s,x) = \infty \quad \text{for } x \neq 0,$$

and hence solving the optimization problems with and without adverse selection requires non-standard considerations.

Portfolio liquidation without adverse selection

In Chapter 2 we define the costs of the strategy u in analogy to Equation (1) by the linear-quadratic functional

$$\mathbb{E}\Big[\int_t^T (\xi(s)^\top \Lambda \xi(s) + \alpha X^u(s)^\top \Sigma X^u(s))ds\Big].$$

The investor aims to minimize these costs among all strategies that fulfill suitable measurability and integrability conditions and the Liquidation Constraint (3).

Because of the singularity of the value function at the terminal time T, we approximate the optimal liquidation problem by a sequence of modified optimization problems with increasing end-costs

$$g(l, X^u(T)) := l \cdot X^u(T)^\top X^u(T) \quad \text{for } l \to \infty.$$

For fixed $l > 0$, a quadratic ansatz yields a candidate w for the solution of the corresponding HJB equation. It turns out that w takes the form

$$w(l,t,x) = x^\top C(l,t)x \quad \text{for } t \in [0,T],\ x \in \mathbb{R}^n, \tag{4}$$

where the matrix-valued function $C(l,\cdot)$ is the solution of the matrix initial value problem

$$\begin{aligned} C' &= C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha\Sigma \\ C(T) &= lI \end{aligned} \tag{5}$$

for

$$\tilde{C} = \operatorname{diag}\left(\frac{\theta_i}{c_{i,i}}\right)$$

(Corollary 2.1.6). In particular, the terminal condition $w(t,x,T) = g(l,x)$ is satisfied. Because of the non-linear term $C^\top \tilde{C} C$, this is not an initial value problem for a Riccati matrix differentiable equation if $n \geq 2$. For $n = 1$,

$$C^\top \tilde{C} C = \theta C,$$

and (5) is an initial value problem for a scalar Riccati differential equation with constant coefficients. The solution for such a problem is well-known in closed form and the limit for $l \to \infty$ exists and can easily be given in closed form. This yields a candidate for the value function of the original optimization problem with liquidation constraint. The fact that all objects are given in closed form reduces the complexity of the optimization problem significantly and yields a much simpler verification argument.

For $n \geq 2$, we first have to prove that the Initial Value Problem (5) admits a solution on the whole interval $[0,T]$. To this end, we establish the matrix inequalities

$$0 \leq C\tilde{C}C \leq \sum_{i=1}^{n} \theta_i C \quad \text{for } C \text{ positive definite}$$

(Theorem 2.2.4 and Corollary 2.2.5). These inequalities enables us to prove existence of the solution of (5) via a well-known comparison theorem for Riccati matrix differential equations and at the same time to obtain explicit upper and lower bounds for this solution (Theorem 2.2.8). The bounds are key. They allow us to prove that w as in Equation (4) is indeed the value function of the optimization problem with finite end-costs (Section 2.3). Furthermore, they yield an existence result for the limit of the solutions of these optimization problems as $l \to \infty$ and thus a candidate for the value function of the optimal liquidation problem (Proposition 2.4.3). Finally, the limits of the bounds are limits for this candidate value function (Theorem 2.4.4), which enables us to prove a verification theorem for the optimal liquidation problem (Theorem 2.4.10).

We analyze the properties of the optimal strategy and the value function in Section 2.5. For $n = 1$, all objects can be derived in closed form and we can prove monotonicity properties that resemble the properties of the discrete-time setting. It is an interesting feature that the risk costs are decreasing in the intensity of the Poisson process θ, while they are increasing in the probability of execution p for small p in the corresponding discrete-time setting. This effect disappears if we let the number of trading times $N+1$ tend to infinity.

At last, we prove that the solution of the single asset discrete-time optimization problem (with constant step-size) converges to the solution of the continuous-time optimization problem as the number of trading times tends to infinity in Section 2.6. We give heuristic arguments that suggest that an according convergence result also holds for general n.

Single asset liquidation with adverse selection

In Chapter 3 we consider adverse selection in continuous time and analyze the single asset liquidation problem corresponding to the discrete-time Cost Functional (2). In analogy to (2) the costs of a continuous-time trading strategy $u = (\xi, \eta)$ are given by

$$\mathbb{E}\Big[\int_t^T (\Lambda\xi(s)^2 + p\Gamma|\eta(s)| + \alpha\Sigma X^u(s)^2)ds\Big].$$

Again, these costs are to be minimized among all strategies that satisfy the Liquidation Constraint (3) and the value function has a singularity at the terminal time T. This time the optimization problem is not linear-quadratic, and hence the value function is not quadratic. This complicates the analysis significantly.

The key is to find a suitable candidate for the value function, which we can no longer expect to be of linear-quadratic form. Instead, the results of the corresponding discrete-time optimization problem suggest that the value function is a quadratic polynomial for small and for large asset positions. For intermediate positions we have to "interpolate" these polynomials. In order to obtain this "interpolation procedure", we make the ansatz that a candidate w for the value function must have the form

$$w(t,x) = \bar{C}_1(t,x)x^2 + \bar{C}_2(t,x)|x| + \bar{C}_3(t,x),$$

with $\bar{C}_1(t,\cdot)$, $\bar{C}_2(t,\cdot)$ and $\bar{C}_3(t,\cdot)$ constant in x for $|x|$ smaller respectively larger than appropriate time-dependent thresholds. The main step for the derivation of the "coefficients" $\bar{C}_1$, $\bar{C}_2$ and $\bar{C}_3$ is the educated guess that

$$x^2\frac{\partial \bar{C}_1}{\partial x}(t,x) + |x|\frac{\partial \bar{C}_2}{\partial x}(t,x) + \frac{\partial \bar{C}_3}{\partial x}(t,x) = 0 \tag{6}$$

(Equation (3.7)). Under the assumption that (6) holds and after a slight change of notation, the HJB equation yields scalar initial value problems for the coefficients (Equations (3.17) - (3.20)). These are initial value problems for a scalar Riccati differential equation with constant coefficients respectively for linear differential equations and can be solved in closed form. Hence, we obtain a candidate w for the value function in closed form (Equations (3.21) - (3.23)).

We analyze w in detail in Section 3.3. We show in Corollary 3.3.5 that w is well-defined. Subsequently, we prove differentiability of w on $[0,T) \times \mathbb{R}$ and compute the partial derivatives (Theorem 3.3.6). The main step in the proof of Theorem 3.3.6 is to verify that the Assumption (6) is indeed true (Lemma 3.3.9). The proofs use rather elementary methods but are cumbersome and tedious.

The analysis of the candidate value function w enables us to show that it fulfills the corresponding HJB equation with singularity at the terminal time T (Theorem 3.3.11). The unique maximizer for the HJB equation provides a candidate for the optimal liquidation strategy. Both the trading intensity in the primary venue and the orders in the dark pool are affine linear for small and for large asset positions. For intermediate asset positions they are given by “interpolations” of these affine linear functions.

In spite of the singularity at the terminal time T, the fact that the candidate value function is given in closed form allows us to finally prove the desired verification theorem (Theorem 3.4.4). Because of the singularity, this involves taking the limit $s \to T-$, and hence requires preliminary considerations (cf., e.g., Lemma 3.4.5).

Subsequently, we analyze the properties of the optimal liquidation strategy and the value function in Section 3.5. Again, these resemble the properties of the respective objects of the corresponding discrete-time setting.

Acknowledgments

First of all, I would like to thank my advisor Ulrich Horst for his support, his helpful advice and his encouragement concerning all aspects of my thesis and my work as an employee at his institute. I am grateful to Torsten Schöneborn for the fruitful collaboration on dark pools, the invaluable research ideas and also for his patience and constant encouragement. I am thankful to my father for his help with Riccati matrix differential equations and to Christoph Baumgarten, Hartmuth Henkel, Mark Van Achter and Nicholas Westray for helpful discussions and useful advice. I would also like to thank Stefan Ankirchner, Darrell Duffie and Ivar Ekeland for their kindness to be available as co-examiners.

During the work on this thesis, I was a scientific employee at Humboldt-Innovation GmbH in the context of the Quantitative Products Laboratory of Deutsche Bank, Humboldt-Universität zu Berlin and Technische Universität Berlin. I am grateful for the financial support, which enabled me to concentrate on the research for my thesis.

1. Optimal liquidation in discrete time

This chapter considers a discrete-time market model for liquidating a portfolio simultaneously at the primary exchange and in a dark pool. We introduce the model in Section 1.1. It is flexible enough to allow for a very general form of the price impact of the large investor and the liquidity in the dark pool. The objective of the investor is to minimize a cost functional which accounts for both the price impact costs of trading and for risk costs originating from the market risk of the portfolio. This gives rise to a stochastic control problem, which we subsequently solve for different specifications of the model.

In Section 1.2 we make only minor additional assumptions. We assume that the liquidities of the n assets in the dark pool are independent and that the underlying probability space is finite. Despite the general structure, we are able to prove existence and uniqueness of the optimal trading strategy, where uniqueness follows from an extensive non-standard convexity argument.

In order to derive additional properties of the optimal strategy, we consider a more specific model in Section 1.3; we assume the price impact to be linear and temporary and that dark pool orders are executed fully or not at all. The resulting cost functional is of linear-quadratic form, and we are able to derive an explicit backward recursion for the quadratic value function and the linear optimal strategy for *multi asset* liquidation and closed form solutions for *single asset* liquidation.

In Section 1.4 we generalize this setting for *single asset* liquidation by introducing adverse selection. The cost functional is not linear-quadratic and thus the structure of the value function and the optimal strategy are significantly more complicated. We are still able to solve the optimal control problem and derive closed form solutions for the value function and the optimal strategy via extensive heuristic considerations.

We close the chapter by considering a different execution price for orders in the dark pool in Section 1.5. If the price in the dark pool includes the market impact generated in the primary venue, market manipulation strategies can become possible and optimal strategies might not exist. Both effects can be avoided by adapting the parameters carefully, e.g., by assuming that dark pool liquidity is bounded and that adverse selection is sufficiently strong.

1.1. Model description

The market we consider consists of n risky assets. Time is discrete and trades can be executed at the (not necessarily equidistant) time points

$$0 = t_0 < \cdots < t_N = T.$$

For example, the distance can be taken in volume time to adjust for the U-shaped intraday pattern of market volatility and liquidity.

At each time point the large investor as well as a number of noise traders execute orders. We denote the orders of the investor at time t_i at the primary venue by

$$x(t_i) = (x_1(t_i), \ldots, x_n(t_i))^\top \in \mathbb{R}^n$$

and in the dark pool by

$$y(t_i) = (y_1(t_i), \ldots, y_n(t_i))^\top \in \mathbb{R}^n.$$

Positive entries denote sell orders and negative entries denote buy orders.

Uncertainty and the evolution of information in the market is described by a stochastic basis

$$(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_{t_i})_{i=0,\ldots,N}, \tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_{t_i})_{i=0,\ldots,N}, \mathbb{P})$$

with two filtrations $\mathbb{F}$, $\tilde{\mathbb{F}}$, where $\tilde{\mathcal{F}}_{t_i} \subseteq \mathcal{F}_{t_i}$. The larger filtration $\mathbb{F}$ reflects the full information in the market, whereas $\tilde{\mathbb{F}}$ reflects the partial information available to the investor. We give a more precise specification of the two filtrations later.

In the following Sections 1.1.1 and 1.1.2 we describe the different effects of the orders $x(t_i)$ and $y(t_i)$. The execution of the order $x(t_i)$ at the primary venue is *guaranteed* but has an adverse effect on the market price. The execution of the order $y(t_i)$ in the dark pool is *uncertain* but has no price impact (irrespective of whether it is executed or not). In Section 1.1.3 we define the trading objective of the investor and specify the set of admissible strategies.

1.1.1. Transaction price and impact of primary venue orders

We assume that the transaction price $P(t_i) \in \mathbb{R}^n$ at the exchange at time t_i can be decomposed into the price impact of the primary venue trades $(x(t_j))_{j=0,\ldots,i}$ of the large trader and the "fundamental" asset price $\tilde{P}(t_i) \in \mathbb{R}^n$ that would have occurred in the absence of large trades. We model the fundamental asset price $\tilde{P}(t_i)$ as a stochastic process with *independent increments* $\epsilon(t_i) \in \mathbb{R}^n$ adapted to $\tilde{\mathbb{F}}$:

$$\tilde{P}(t_{i+1}) = \tilde{P}(t_i) + \epsilon(t_{i+1}).$$

We do not make assumptions on the distributions of the $\epsilon(t_i)$. In particular, they can have different distributions. The random price changes $\epsilon(t_i)$ reflect the noise traders' actions as well as all external events, e.g., news.

We allow a general form of the impact of the trades $x(t_0)$, …, $x(t_i)$ on the transaction price $P(t_i)$:

$$P(t_i) = \underbrace{\tilde{P}(t_i)}_{\substack{\uparrow \\ \text{“Fundamental”} \\ \text{asset price}}} - \underbrace{f_i(x(t_0), \ldots, x(t_i))}_{\substack{\uparrow \\ \text{Price impact} \\ \text{of seller}}},$$

where $f_i : \mathbb{R}^{n\times(i+1)} \to \mathbb{R}^n$. By allowing f_i to depend on $x(t_i)$, we allow the order $x(t_i)$ to influence its own execution price (in the form of a temporary price impact). We define the *price impact costs of trading* as

$$\sum_{i=0}^{N} x(t_i)^\top f_i(x(t_0), \ldots, x(t_i)).$$

Assumption 1.1.1. *The price impacts fulfill the following two conditions.*

(i) $\sum_{i=0}^{N} x(t_i)^\top f_i(x(t_0), \ldots, x(t_i))$ *is strictly convex in* $(x(t_0), \ldots, x(t_N))$.

(ii) $\sum_{i=0}^{N} x(t_i)^\top f_i(x(t_0), \ldots, x(t_i))$ *grows superlinearly, i.e.,*

$$\lim_{\|(x(t_0),\ldots,x(t_N))\|\to\infty} \frac{\sum_{i=0}^{N} x(t_i)^\top f_i(x(t_0), \ldots, x(t_i))}{\|(x(t_0), \ldots, x(t_N))\|} = \infty.$$

This framework generalizes many of the existing market impact models of liquidity.

Examples 1.1.2. *(i) The model suggested by Almgren and Chriss [2001] is equivalent to assuming that the* $\epsilon(t_i)$ *are identically normally distributed and*

$$f_i(x(t_0), \ldots, x(t_i)) = \mathrm{TempImp}(x(t_i)) + \sum_{j=0}^{i-1} \mathrm{PermImp}(x(t_j)).$$

Here, $\mathrm{PermImp}, \mathrm{TempImp} : \mathbb{R}^n \to \mathbb{R}^n$ *are functions describing the permanent and temporary market impact of a trade. If these functions are linear, we have*

$$f_i(x(t_0), \ldots, x(t_i)) = \Lambda_i x(t_i) + \sum_{j=0}^{i-1} \Gamma_{i,j} x(t_j)$$

for matrices $\Lambda_i, \Gamma_{i,j} \in \mathbb{R}^{n\times n}$ *(*$i = 0, \ldots, N$*,* $j = 0, \ldots, i-1$*). In terms of the matrix*

$$\tilde{\Lambda} := \begin{pmatrix} \Lambda_0 & \frac{1}{2}\Gamma_{1,0}^\top & \cdots & \frac{1}{2}\Gamma_{N,0}^\top \\ \frac{1}{2}\Gamma_{1,0} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{1}{2}\Gamma_{N,N-1}^\top \\ \frac{1}{2}\Gamma_{N,0} & \cdots & \frac{1}{2}\Gamma_{N,N-1} & \Lambda_N \end{pmatrix}$$

we get

$$\sum_{i=0}^{N} x(t_i)^\top f_i(x(t_0), \ldots, x(t_i)) = \tilde{x}^\top \tilde{\Lambda} \tilde{x}$$

for $\tilde{x} = (x(t_0), \dots, x(t_N))^\top \in \mathbb{R}^{n(N+1)}$*. Thus, Assumption 1.1.1 is fulfilled if the* Λ_i *are symmetric and* $\tilde{\Lambda}$ *is positive definite.*

(ii) The limit order book model introduced by Obizhaeva and Wang [2006] is included in our model as a single asset example if we again assume that the $\epsilon(t_i)$ *are identically normally distributed and that the price impact is given by*

$$f_i(x(t_0), \dots, x(t_i)) = \gamma\Big(\sum_{j=0}^{i-1} x(t_j) + \frac{x(t_i)}{2}\Big) + \lambda\Big(\sum_{j=0}^{i-1} e^{-\rho(t_i - t_j)} x(t_j) + \frac{x(t_i)}{2}\Big)$$

for constants $\gamma \geq 0$*,* $\lambda > 0$ *and* $\rho > 0$*. It is elementary but somewhat tedious to show that Assumption 1.1.1 (i) is satisfied. We omit this proof here.*

1.1.2. Trade execution in the dark pool

Contrary to the primary venue, the dark pool does not guarantee trade execution since it only provides limited liquidity. We introduce the $\mathcal{F}_{t_{i+1}}$-measurable random variables

$$a(t_i), b(t_i) \in [0, \infty]^n$$

that model the liquidity which can be drawn upon in the time-interval $[t_i, t_{i+1})$ for buy (ask side) and sell (bid side) orders, respectively. The dark pool liquidity is in general *unknown* to the investor and therefore in general *not* $\tilde{\mathcal{F}}_{t_{i+1}}$-measurable.

Assumption 1.1.3. *The dark pool liquidity fulfills the following three conditions.*

(i) For $i = 0, \dots, N$*,* $a(t_i)$ *and* $b(t_i)$ *are independent of previous liquidity in the dark pool* $a(t_0), \dots, a(t_{i-1}), b(t_0), \dots, b(t_{i-1})$ *and of previous price moves* $\epsilon(t_1), \dots, \epsilon(t_i)$*.*

(ii) For $i = 0, \dots, N$*,* $k = 1, \dots, n$*,*

$$\mathbb{P}[a_k(t_i) = 0] > 0, \quad \mathbb{P}[b_k(t_i) = 0] > 0.$$

(iii) For all $i = 0, \dots, N$*,* $k = 1, \dots, n$*,* $p > q \geq 0$ *with*

$$\mathbb{P}[a_k(t_i) = p], \mathbb{P}[a_k(t_i) = q] > 0 \quad \textit{respectively} \quad \mathbb{P}[b_k(t_i) = p], \mathbb{P}[b_k(t_i) = q] > 0,$$

we have

$$0 \geq \mathbb{E}[\epsilon_k(t_{i+1}) | a_k(t_i) = q] \geq \mathbb{E}[\epsilon_k(t_{i+1}) | a_k(t_i) = p], \tag{1.1}$$
$$0 \leq \mathbb{E}[\epsilon_k(t_{i+1}) | b_k(t_i) = q] \leq \mathbb{E}[\epsilon_k(t_{i+1}) | b_k(t_i) = p]. \tag{1.2}$$

Assumption 1.1.3 allows for a dependence of the liquidity parameters $a(t_i)$ and $b(t_i)$ and the price move $\epsilon(t_{i+1})$. This enables us to incorporate the simultaneous occurrence of price jumps and liquidity in the dark pool which can lead to *adverse selection* (see Section 1.4). Assumption 1.1.3 (iii) is needed in order to ensure uniqueness of the optimal strategy. Economically, it means that price moves in the market are monotone with respect to dark pool liquidity, i.e., the stronger the demand in the dark pool, the stronger the price at the exchange is expected to move upwards, and the stronger the supply in the dark pool, the stronger the price is expected to move downwards. In other words: a large amount of liquidity in the dark pool could be a sign for an impending favorable price move. The case of strict inequality in Inequalities (1.1) or (1.2) for some $p > q \geq 0$ can lead to adverse selection and is studied in detail in Section 1.4.

The amount
$$z(t_i) = (z_1(t_i), \dots, z_n(t_i))^\top \in \mathbb{R}^n$$
which is executed in the dark pool between time t_i and t_{i+1} is given by

$$z_k(t_i) = \begin{cases} \min(y_k(t_i), b_k(t_i)) & \text{if } y_k(t_i) \geq 0 \\ -\min(-y_k(t_i), a_k(t_i)) & \text{if } y_k(t_i) < 0, \end{cases}$$

where $y_k(t_i)$ is the order in the dark pool for the k^{th} asset at time t_i. In contrast to the dark pool liquidity $(a(t_i), b(t_i))$, which is general *not* known to the investor at the end of the trading period $[t_i, t_{i+1})$, $z(t_i)$ is *known* to the investor at time t_{i+1}, and we assume that it is $\tilde{\mathcal{F}}_{t_{i+1}}$-measurable. Having introduced all random objects, we can hence interpret the two filtrations as follows. For $i = 0, \dots, N$,

$$\begin{aligned} \mathcal{F}_{t_i} &= \sigma\big(\epsilon(t_j), a(t_{j-1}), b(t_{j-1});\ j \leq i\big), \\ \tilde{\mathcal{F}}_{t_i} &= \sigma\big(\epsilon(t_j), z(t_{j-1});\ j \leq i\big). \end{aligned}$$

Mathematically, this distinction does not play a role by Assumption 1.1.3 (i).

While the dark pool has no impact on prices at the primary exchange, it is less clear to which extent the price impact f_i of the exchange is reflected in the transaction price of the dark pool. If, for example, the price impact f_i is realized predominantly in the form of a widening spread, then the impact on dark pools that monitor the mid quote can be much smaller than f_i. In Sections 1.2 - 1.4 we make the simplifying assumption that trades in the dark pool are not influenced by the price impact f_i at all, i.e., that they are executed at the fundamental price $\tilde{P}(t_i)$. If trading in the dark pool reflects the price impact f_i, then market manipulating strategies can become profitable. We investigate this phenomenon in Section 1.5.

1.1.3. Liquidation problem

For fixed $i = 0, \dots, N$, we consider an investor who has executed trades
$$x(t_0), \dots, x(t_{i-1}) \in \mathbb{R}^n$$
at times $t_0, \dots, t_{i-1}$ and needs to liquidate a portfolio
$$X(t_i) = (X_1(t_i), \dots, X_n(t_i))^\top \in \mathbb{R}^n$$

of risky assets within a finite time-horizon $[t_i, T]$. For $X_k(t_i) > 0$, this implies liquidating a long position in asset k (selling), whereas $X_k(t_i) < 0$ implies liquidating a short position in asset k (buying). In both cases, we speak of "liquidation" or "sale". We require that at all times $t_j \geq t_i$ the investor's orders $x(t_j)$ and $y(t_j)$ are adapted to $\tilde{\mathcal{F}}_{t_j}$ and therefore only depend on past price-moves $\epsilon(t_1), \dots, \epsilon(t_j)$ and executed orders $z(t_0), \dots, z(t_{j-1})$. This includes deterministic (also called *static*) strategies, which do not depend on *any* of the quantities $\epsilon(t_l)$ or $z(t_l)$.

Definition 1.1.4. *Let $i = 0, \dots, N$ and $X(t_i) \in \mathbb{R}^n$ be the portfolio position at time t_i. We call a sequence of $\tilde{\mathbb{F}}$-adapted orders*

$$(x, y) = (x(t_j), y(t_j))_{j=i,\dots,N}$$

an admissible liquidation strategy *if it fulfills the following conditions.*

(i)

$$\sum_{j=i}^{N} \big(x(t_j) + z(t_j)\big) = X(t_i) \quad a.s.$$

(ii) For $j = i, \dots, N$, $k = 1, \dots, n$,

$$\mathbb{P}\big[-a_k(t_j) \leq y_k(t_j) \leq b_k(t_j)\big] > 0.$$

We denote the set of admissible liquidation strategies by $\mathbb{A}(t_i, X(t_i))$.

Let us shortly comment on Definition 1.1.4. We recursively define for $j \geq i + 1$,

$$X(t_{j+1}) := X(t_j) - x(t_j) - z(t_j). \tag{1.3}$$

By abuse of notation, the liquidation constraint in (i) is then equivalent to

$$X(t_{N+1}) = 0 \quad \text{a.s.}$$

By Assumption 1.1.3 (ii), all admissible liquidation strategies must satisfy

$$x(t_N) = X(t_N), \quad y(t_N) = z(t_N) = 0 \quad \text{a.s.}$$

Definition 1.1.4 (ii) is required in order to ensure uniqueness of the optimal strategy. Note that we assume implicitly that the essential maxima of $a_k(t_j)$ and $b_k(t_j)$ are known to the investor. Due to order submission fees, it is natural in practice to assume that an investor only submits orders that have a positive probability of complete execution; orders not satisfying Definition 1.1.4 (ii) cannot generate additional value (compared to strategies that do satisfy Definition 1.1.4 (ii)).

The following definition is based on Perold [1988] and serves as the investor's performance measure.

Definition 1.1.5. *For $i = 0, \dots, N$, past trades $x(t_0), \dots, x(t_{i-1}) \in \mathbb{R}^n$, portfolio position $X(t_i) \in \mathbb{R}^n$ at time t_i and an admissible liquidation strategy $(x, y) \in \mathbb{A}(t_i, X(t_i))$, the* implementation shortfall *is given by*

$$\begin{aligned}\mathcal{R}(t_i) &:= \mathcal{R}(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i); (x, y)) \\ &:= X(t_i)^\top \tilde{P}(t_i) - \sum_{j=i}^{N} \left(x(t_j)^\top P(t_j) + z(t_j)^\top \tilde{P}(t_j)\right) \\ &= \sum_{j=i}^{N} \Big(x(t_j)^\top \Big(\tilde{P}(t_i) - \tilde{P}(t_j) + f_j(x(t_0), \dots, x(t_j))\Big) + z(t_j)^\top (\tilde{P}(t_i) - \tilde{P}(t_j))\Big).\end{aligned}$$

The trade-off between expected proceeds and risk is an important driver of optimal liquidation and has been the focus of several investigations including Almgren and Chriss [2001], Almgren and Lorenz [2007], Schied and Schöneborn [2009] and Schied et al. [2010]. Here, we assume that the investor wants to minimize the following function of execution costs:

$$\begin{aligned}&J(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i); (x, y)) \\ &\quad := \mathbb{E}\Big[\mathcal{R}(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i); (x, y))\Big] + \alpha \mathbb{E}\Big[\sum_{j=i}^{N} X(t_j)^\top \Sigma(t_{j+1}) X(t_j)\Big],\end{aligned}$$

where $\alpha \geq 0$ is the coefficient of risk-aversion and $\Sigma(t_j)$ is the covariance matrix of the increments $\epsilon_k(t_j)$. The risk costs

$$\alpha \sum_{j=i}^{N} X(t_j)^\top \Sigma(t_{j+1}) X(t_j)$$

reflect the market risk of the portfolio and thus penalize slow execution and poorly diversified portfolios. In the setting of Almgren and Chriss [2001] for optimal liquidation without dark pools, this is equivalent to minimizing the mean-variance functional

$$\mathbb{E}[\mathcal{R}(t_i)] + \alpha \operatorname{Var}[\mathcal{R}(t_i)]$$

over all deterministic strategies. Schied et al. [2010] show that this in turn is equivalent to maximizing the utility of investors with constant absolute risk-aversion. The value function of the optimization problem is thus given by

$$\begin{aligned}&v(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i)) \\ &\quad := \inf_{(x,y) \in \mathbb{A}(t_i, X(t_i))} J(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i); (x, y)). \qquad (\mathrm{OPT}^{\mathrm{gen}}_{\mathrm{dis}})\end{aligned}$$

We call an admissible liquidation strategy $(x, y) \in \mathbb{A}(t_i, X(t_i))$ *optimal* if it realizes the minimum in Equation ($\mathrm{OPT}^{\mathrm{gen}}_{\mathrm{dis}}$)and denote optimal strategies by

$$(x^*, y^*)$$

in the remainder of the chapter. The amount executed in the dark pool in $[t_i, t_{i+1})$ associated with the optimal order $y^*(t_i)$ is denoted by

$$z^*(t_i).$$

1.2. Optimal liquidation

In this section we state sufficient conditions for the Optimization Problem ($\mathrm{OPT}^{\mathrm{gen}}_{\mathrm{dis}}$) to admit a unique solution. To this end, we require the following additional assumptions.

Assumption 1.2.1. *(i) The space Ω is finite and $\mathbb{P}[\omega] > 0$ for all $\omega \in \Omega$.*

(ii) For $i = 0, \dots, N-1$,

$$(a_1(t_i), b_1(t_i)), \dots, (a_n(t_i), b_n(t_i))$$

are independent.

The results of this section extend to infinite state space and possibly dependent dark pool liquidities of the n assets if the price increments ϵ, the liquidity variables (a, b) and the price impact functions f_i satisfy suitable conditions (cf. Sections 1.3 and 1.4). The additional assumptions enable us to prove the main result of this section.

Theorem 1.2.2. *Let $i = 0, \dots, N-1$, $x(t_0), \dots, x(t_{i-1}) \in \mathbb{R}^n$ be the previous trades of the investor and $X(t_i) \in \mathbb{R}^n$ be the portfolio position at time t_i. Assume that the processes ϵ, a and b and the functions f_j fulfill the assumptions of Section 1.1 and Assumption 1.2.1.*

Then there exists a unique optimal strategy $(x^*, y^*) \in \mathbb{A}(t_i, X(t_i))$ realizing the minimum in Equation ($\mathrm{OPT}^{\mathrm{gen}}_{\mathrm{dis}}$).

The remainder of the section is devoted to the proof of Theorem 1.2.2. In Section 1.2.1 we prove existence of optimal liquidation strategies (Proposition 1.2.3). In Section 1.2.2 we show that the function

$$\begin{aligned} &H_i(x(t_0), \dots, x(t_{i-1}), X(t_i)) \\ &\quad := v(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i)) + \sum_{j=0}^{i-1} x(t_j)^\top f_j(x(t_0), \dots, x(t_j)) \end{aligned} \tag{1.4}$$

is strictly convex in $(x(t_0), \dots, x(t_i), X(t_i))$, from which we deduce uniqueness of the optimal strategy (Theorem 1.2.4).

1.2.1. Existence of optimal trading strategies

We first prove existence of optimal trading strategies. We only require (i) of the additional Assumption 1.2.1 for the proof. Assumption 1.2.1 (ii) is only needed for the proof of uniqueness in Section 1.2.2.

Proposition 1.2.3. *Let $i = 0, \ldots, N-1$, $x(t_0), \ldots, x(t_{i-1}) \in \mathbb{R}^n$ be the previous trades of the investors and $X(t_i) \in \mathbb{R}^n$ be the portfolio position at time t_i. Assume that the processes ϵ, a and b and the functions f_j fulfill the assumptions of Section 1.1 and Assumption 1.2.1 (i).*

Then there exists an optimal strategy $(x^*, y^*) \in \mathbb{A}(t_i, X(t_i))$ realizing the minimum in Equation (OPT$^{\text{gen}}_{\text{dis}}$).

Proof. Instead of describing a strategy $(x, y) \in \mathbb{A}(t_i, X(t_i))$ as a stochastic process, we can alternatively describe it as a vector. Let therefore

$$\Omega = \{\omega_1, \ldots, \omega_M\}.$$

By abuse of notation, we write

$$w = \Big(\underbrace{x(t_i,\omega_1),\ldots,x(t_i,\omega_M),x(t_{i+1},\omega_1),\ldots,x(t_{i+1},\omega_M),\ldots,x(t_N,\omega_1),\ldots,x(t_N,\omega_M)}_{=:w_x},$$
$$\underbrace{y(t_i,\omega_1),\ldots,y(t_N,\omega_M)}_{=:w_y}\Big)^\top \in \mathbb{R}^{n\times 2\times M\times(N+1-i)}.$$

The objective function

$$\mathcal{C}(w) := J(t_i, x(t_0), \ldots, x(t_{i-1}), X(t_i); w)$$

is continuous in the strategy $w \in \mathbb{R}^{n\times 2\times M\times(N+1-i)}$, and the set of admissible strategies corresponds to a closed subset of $\mathbb{R}^{n\times 2\times M\times(N+1-i)}$. We show that

$$\lim_{\|w\|\to\infty} \mathcal{C}(w) = \infty, \tag{1.5}$$

where $\|\cdot\|$ is the maximum norm on $\mathbb{R}^{n\times 2\times M\times(N+1-i)}$. This allows us to restrict w to a bounded set, and so the existence of an optimal strategy follows from continuity of $\mathcal{C}$.

It is sufficient to prove Equation (1.5) for $\alpha = 0$, for which we obtain

$$\begin{aligned}\mathcal{C}(w) &= \mathbb{E}\Big[\sum_{j=i}^N x(t_j)^\top f_j(x(t_0),\ldots,x(t_j))\Big] + \mathbb{E}\Big[\sum_{j=i}^N x(t_j)^\top(\tilde{P}(t_i)-\tilde{P}(t_j))\Big]\\ &\quad + \mathbb{E}\Big[\sum_{j=i}^N z(t_j)^\top(\tilde{P}(t_i)-\tilde{P}(t_j))\Big]\\ &=: \mathcal{C}_1(w)+\mathcal{C}_2(w)+\mathcal{C}_3(w).\end{aligned}$$

Note that $\mathcal{C}_2$ and $\mathcal{C}_3$ are not necessarily bounded from below and that $\mathcal{C}_1(w) > 0$ for $\|w_x\|$ large enough. Therefore, we have to show that $\mathcal{C}_1$ grows faster in w than $|\mathcal{C}_2|$ and $|\mathcal{C}_3|$.

It follows from Assumption 1.1.1 (ii) that

$$\lim_{\|w_x\| \to \infty} \frac{\mathcal{C}_1(w)}{\|w_x\|} = \infty. \tag{1.6}$$

Since Ω is finite, the price process $\tilde{P}$ is bounded, and thus there exists a constant $\tilde{C}$ such that for all $w_x \neq 0$,

$$\begin{aligned} \frac{|\mathcal{C}_2(w)|}{\|w_x\|} &\leq \frac{1}{\|w_x\|} \sum_{j=i}^{N} \mathbb{E}[|x(t_j)^\top (\tilde{P}(t_i) - \tilde{P}(t_j))|] \\ &\leq \frac{\tilde{C}}{\|w_x\|} \sum_{j=i}^{N} \sum_{l=1}^{M} \sum_{k=1}^{n} |x_k(t_j, \omega_l)| \mathbb{P}[\omega_l] \\ &\leq \tilde{C} \cdot n \cdot (N + 1 - i). \end{aligned} \tag{1.7}$$

If the liquidity in the dark pool is bounded for all assets, then $\|w_y\|$ and $|\mathcal{C}_3(w)|$ are bounded and thus

$$\lim_{\|w\| \to \infty} \frac{|\mathcal{C}_3(w)|}{\|w\|} = 0.$$

If not, we obtain similarly as before for $w_y \neq 0$ (recall that $|z| \leq |y|$),

$$\frac{|\mathcal{C}_3(w)|}{\|w_y\|} \leq \tilde{C} \cdot n \cdot (N + 1 - i). \tag{1.8}$$

Finally, a large order in the dark pool at a given point in time requires large orders in the primary venue with positive probability since by Definition 1.1.4 (ii), full execution of the dark pool order is possible, while on the other hand future dark pool orders have positive probability of non-execution (cf. Assumption 1.1.3 (ii)). Thus, there exists a constant C such that

$$\lim_{\|w_y\| \to \infty} \frac{\|w_x\|}{\|w_y\|} > C. \tag{1.9}$$

Equation (1.5) now follows directly from (1.6), (1.7), (1.8) and (1.9). □

1.2.2. Uniqueness of the optimal trading strategy

In order to prove uniqueness of the optimal strategy, we require strict convexity of the function H_i as in Equation (1.4). In Sections 1.3 and 1.4 we consider more specific distributions of the random variables $a(t_i)$ and $b(t_i)$ and obtain a (pathwise) linear dependence of $z(t_i)$ on $y(t_i)$. In that case, using dynamic programming, strict convexity of H_i follows by a standard backward induction (cf., e.g., the proof of Theorem 1.4.2). In general, the dependence of $z(t_i)$ on $y(t_i)$ is non-linear. Therefore, we require non-standard arguments to prove convexity, and this proof is divided into several steps.

Theorem 1.2.4. *Let $i = 0, \dots, N-1$, $x(t_0), \dots, x(t_{i-1}) \in \mathbb{R}^n$ be the previous trades of the investor and $X(t_i) \in \mathbb{R}^n$ be the portfolio position at time t_i. Assume that the processes ϵ, a and b and the functions f_j fulfill the assumptions of Section 1.1 and Assumption 1.2.1. Then:*

(i) The function H_i as in Equation (1.4) *is strictly convex in* $(x(t_0), \dots, x(t_{i-1}), X(t_i))^\top \in \mathbb{R}^{n\times(i+1)}$.

(ii) The optimal trading strategy $(x^*, y^*) \in \mathbb{A}(t_i, X(t_i))$ realizing the minimum in Equation $(\mathrm{OPT}^{\mathrm{gen}}_{\mathrm{dis}})$ *is* unique.

Proof. We proceed by backward induction on i. For $i = N$, the validity of the theorem follows since the only admissible strategy is $x(t_N) = X(t_N)$, $y(t_N) = 0$ due to Assumption 1.1.3 (ii), and the convexity of H_N defined by Equation (1.4) follows directly from the convexity of the price impact cost of trading (Assumption 1.1.1 (i)) as

$$\begin{aligned} H_N(x(t_0), \dots, x(t_{N-1}), X(t_N)) = {} & X(t_N)^\top f_N(x(t_0), \dots, x(t_N)) + \alpha X(t_N)^\top \Sigma(t_{N+1}) X(t_N) \\ & + \sum_{j=0}^{N-1} x(t_j)^\top f_j(x(t_0), \dots, x(t_j)). \end{aligned}$$

For the *induction step* we consider two points

$$(x(t_0), \dots, x(t_{i-1}), X(t_i)) \quad \text{and} \quad (\tilde{x}(t_0), \dots, \tilde{x}(t_{i-1}), \tilde{X}(t_i)).$$

For these points, optimal orders

$$(x(t_i), y(t_i)) := (x^*(t_i), y^*(t_i)) \quad \text{respectively} \quad (\tilde{x}(t_i), \tilde{y}(t_i)) := (\tilde{x}^*(t_i), \tilde{y}^*(t_i))$$

exist by Proposition 1.2.3. We define continuous functions

$$x(t_j, \cdot), y(t_j, \cdot), X(t_i, \cdot) : [0,1] \longrightarrow \mathbb{R}^n, \quad 0 \le j \le i,$$

such that

$$\begin{aligned} & x(t_j, 0) = x(t_j), \quad y(t_j, 0) = y(t_j), \quad X(t_i, 0) = X(t_i), \\ & x(t_j, 1) = \tilde{x}(t_j), \quad y(t_j, 1) = \tilde{y}(t_j), \quad X(t_i, 1) = \tilde{X}(t_i), \quad 0 \le j \le i. \end{aligned}$$

Then by the dynamic programming principle,

$$\begin{aligned} & v(t_i, x(t_0, s), \dots, x(t_{i-1}, s), X(t_i, s)) \\ & \quad \le x(t_i, s)^\top f_i(x(t_0, s), \dots, x(t_i, s)) + \alpha X(t_i, s)^\top \Sigma(t_{i+1}) X(t_i, s) \\ & \qquad + (X(t_i, s) - x(t_i, s))^\top \mathbb{E}[\tilde{P}(t_i) - \tilde{P}(t_{i+1})] - \mathbb{E}[z(t_i, s)^\top (\tilde{P}(t_i) - \tilde{P}(t_{i+1}))] \\ & \qquad + \mathbb{E}[v(t_{i+1}, x(t_0, s), \dots, x(t_i, s), X(t_i, s) - x(t_i, s) - z(t_i, s))] \qquad (1.10) \\ & \quad =: h_i(s), \end{aligned}$$

where Inequality (1.10) is an equality for $s = 0$ and $s = 1$. We now assume that the theorem holds for $i + 1$ and divide the proof of the induction step into three parts.

(i) Let
$$\tilde{h}_i(s) := h_i(s) + \sum_{j=0}^{i-1} x(t_j, s)^\top f_j(x(t_0, s), \ldots, x(t_j, s)).$$
We show that if $\tilde{h}_i$ is strictly convex, then H_i is strictly convex and the optimal strategy at time t_i is unique.

(ii) We define the functions $x(t_j, \cdot), y(t_j, \cdot), X(t_i, \cdot)$ in such a way that $x(t_j, \cdot), \mathbb{E}[z(t_j, \cdot)]$ and $X(t_i, \cdot)$ are affine linear; here, we use the shorthand notation $z(t_i, s) := z(t_i, y(t_i, s))$. This is needed to carry out step (iii).

(iii) We show that if H_{i+1} is strictly convex (induction hypothesis!),
$$(x(t_0), \ldots, x(t_{i-1}), X(t_i), x(t_i), y(t_i)) \neq (\tilde{x}(t_0), \ldots, \tilde{x}(t_{i-1}), \tilde{X}(t_i), \tilde{x}(t_i), \tilde{y}(t_i))$$
and $x(t_j, \cdot), y(t_j, \cdot), X(t_i, \cdot)$ are defined as in (ii), then $\tilde{h}_i$ is strictly convex in s on $[0, 1]$. Hence by (i), H_i is strictly convex.

The natural order of the three steps is (ii), (iii), (i). We start by proving (i) in order to motivate the necessity of the steps (ii) and (iii). The second part is the key step in the proof; the proof of the third part is rather extensive.

(i) Let
$$(x(t_0), \ldots, x(t_{i-1}), X(t_i)) \neq (\tilde{x}(t_0), \ldots, \tilde{x}(t_{i-1}), \tilde{X}(t_i))$$
and $s \in (0, 1)$. Then
$$\begin{aligned}
&H_i((1-s)x(t_0) + s\tilde{x}(t_0), \ldots, (1-s)x(t_{i-1}) + s\tilde{x}(t_{i-1}), (1-s)X(t_i) + s\tilde{X}(t_i)) \\
&\overset{(1.10)}{\leq} \tilde{h}(s) \\
&= \tilde{h}_i((1-s) \cdot 0 + s \cdot 1) \\
&< (1-s)\tilde{h}_i(0) + s\tilde{h}_i(1) \\
&= (1-s)H_i(x(t_0), \ldots, x(t_{i-1}), X(t_i)) + sH_i(\tilde{x}(t_0), \ldots, \tilde{x}(t_{i-1}), \tilde{X}(t_i)),
\end{aligned}$$
where the last equation follows from the fact that we have equality in Inequality (1.10) for $s = 0$ and $s = 1$. Thus, H_i is strictly convex.

For the uniqueness of the optimal strategy, let
$$(x(t_0), \ldots, x(t_{i-1}), X(t_i)) = (\tilde{x}(t_0), \ldots, \tilde{x}(t_{i-1}), \tilde{X}(t_i)).$$
If
$$(x^*(t_i), y^*(t_i)) = (x(t_i), y(t_i)) \neq (\tilde{x}^*(t_i), \tilde{y}^*(t_i)) = (\tilde{x}(t_i), \tilde{y}(t_i)), \tag{1.11}$$

then $\tilde{h}_i(s)$ is strictly convex. On the other hand we have (cf. Inequality (1.10))

$$\tilde{h}_i(0) = \tilde{h}_i(1) \leq \tilde{h}_i(s) \quad \text{for all } s \in (0,1),$$

thus $\tilde{h}_i$ cannot be strictly convex, and (1.11) yields a contradiction.

(ii) We define the functions $x(t_j, \cdot)$ and $X(t_i, \cdot)$ by the convex combinations

$$\begin{aligned} x(t_j, s) &:= (1-s)x(t_j) + s\tilde{x}(t_j) \quad \text{for all } 0 \leq j \leq i,\ 0 \leq s \leq 1, \\ X(t_i, s) &:= (1-s)X(t_i) + s\tilde{X}(t_i) \quad \text{for all } 0 \leq s \leq 1. \end{aligned}$$

Note that if we define $y(t_i, s)$ accordingly, the linearity of $y(t_i, \cdot)$ neither carries over to $z(t_i, s)$ nor to $\mathbb{E}[z(t_i, s)]$. The key step in the proof is to define $y(t_i, s)$ in such a way that

$$s \mapsto \mathbb{E}[z(t_i, s)]$$

is affine linear. We set $y(t_i, s)$ such that

$$\mathbb{E}[z(t_i, s)] = (1-s)\mathbb{E}[z(t_i)] + s\mathbb{E}[\tilde{z}(t_i)].$$

To this end, we define the function

$$g(y) := \mathbb{E}[z(y)].$$

It is injective by Definition 1.1.4 (ii) and continuous. We define

$$y(t_i, s) := g^{-1}\big((1-s)\mathbb{E}[z(t_i)] + s\mathbb{E}[\tilde{z}(t_i)]\big),$$

in particular $y(t_i, 0) = y^*(t_i)$, $y(t_i, 1) = \tilde{y}^*(t_i)$ and $y(t_i, \cdot)$ is continuous on $[0,1]$. Note that $y(t_i, \cdot)$ is piecewise affine linear but not affine linear in general; also, $\mathbb{E}[z(t_i, \cdot)]$ is affine linear but pathwise $z(t_i, \cdot)$ is *not* affine linear.

(iii) If $z(t_i, \cdot)$ was pathwise affine linear on the whole interval $[0,1]$, then strict convexity of H_{i+1} would directly transfer to strict convexity of $\tilde{h}_i$ via Inequality (1.10). As we only have linearity of $\mathbb{E}[z(t_i, \cdot)]$, the argumentation becomes rather tedious and extensive.

The only points where convexity of $\tilde{h}_i$ can break down are the points s_j at which the slope changes for $y_k(t_i, \cdot)$ for some $k = 1, \ldots, n$ (cf. Figure 1.1).

We denote the finitely many points at which there is a coordinate k such that

$$\mathbb{P}[-y_k(t_i, s_j) = a_k(t_i)] > 0 \quad \text{or} \quad \mathbb{P}[y_k(t_i, s_j) = b_k(t_i)] > 0 \tag{1.12}$$

by $0 < s_1 < \cdots < s_{M'} < 1$. We can assume without loss of generality that at each s_j there is exactly one coordinate k_j such that (1.12) holds. In the case that there are multiple such k's, an arbitrary small perturbation of $a(t_i)$ and $b(t_i)$ removes multiplicity, and a simple approximation argument using the fact that

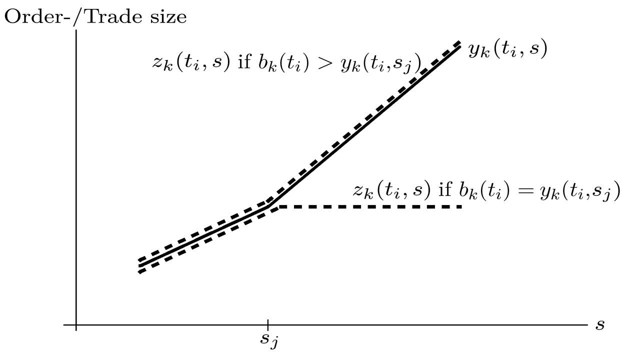


Figure 1.1.: If $\mathbb{P}[y_k(t_i, s_j) = b_k(t_i)] > 0$ for an order $y_k(t_i, s_j) > 0$, then the slope of $y_k(t_i, \cdot)$ can change at s_j, and $z_k(t_i, \cdot)$ is *not* pathwise affine linear.

$z(t_i)$ is continuous in $a(t_i)$ and $b(t_i)$ extends the desired result to full generality.

On (s_j, s_{j+1}), the strict convexity of $\tilde{h}_i$ is clear since the map

$$s \mapsto (x(t_0, s), \ldots, x(t_i, s), z(t_i, s), X(t_i, s)) \tag{1.13}$$

is *pathwise affine linear* (note that $y_k(t_i, \cdot)$ is affine linear on (s_j, s_{j+1})), and thus $\tilde{h}_i$ is strictly convex by the strict convexity of H_{i+1} via Equation (1.10).

Let therefore $j = 1, \ldots, M'$ and $k \in \{1, \ldots, n\}$ such that

$$\mathbb{P}\big[\{y_k(t_i, s_j) = -a_k(t_i)\} \cup \{y_k(t_i, s_j) = b_k(t_i)\}\big] > 0.$$

We first consider the case

$$y_k(t_i, s_j) > 0, \quad \text{i.e.,} \quad \mathbb{P}\big[y_k(t_i, s_j) = b_k(t_i)\big] > 0.$$

Then $y_k(t_i, s) > 0$ for all $s \in (s_{j-1}, s_{j+1})$ by Assumption 1.1.3 (ii). We assume without loss of generality that $y_k(t_i, s_{j-1}) < y_k(t_i, s_{j+1})$ (the proof for the case $y_k(t_i, s_{j-1}) = y_k(t_i, s_{j+1})$ is straightforward) and define

$$A := \{y_k(t_i, s_j) \le b_k(t_i)\} \subseteq \Omega.$$

Let

$$\bar{z}_k(t_i, s) := \begin{cases} \mathbb{E}\big[z_k(t_i, s)|A\big] & \text{on } A \\ z_k(t_i, s) & \text{otherwise.} \end{cases} \tag{1.14}$$

$$\bar{z}_l(t_i, s) := z_l(t_i, s) \quad \text{for } l \neq k.$$

Note that for $l \neq k$, $z_l(t_i, s)$ is *pathwise affine linear* on the whole interval (s_{j-1}, s_{j+1}). On $\Omega \setminus A$, $z_k(t_i, s)$ is independent of $s \in (s_{j-1}, s_{j+1})$. On A, we

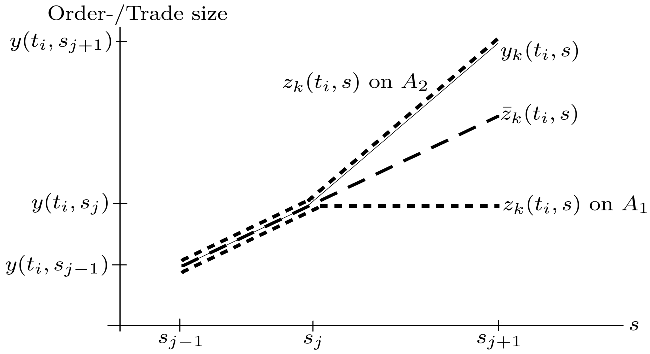


Figure 1.2.: Orders $y_k(t_i, s)$ and associated trades $z_k(t_i, s)$ on the interval (s_{j-1}, s_{j+1}). The thin solid line represents the orders $y_k(t_i, s)$. Note that the slope of $y_k(t_i, s)$ is larger for $s > s_j$ than for $s < s_j$. The dotted lines represent the trades $z_k(t_i, s)$ on A_1 and A_2, respectively. The affine linear function $\bar{z}(t_i, s)$ defined in Equation (1.14) is represented by the dashed line.

have

$$\mathbb{E}[z_k(t_i, s)|A] = \frac{1}{\mathbb{P}(A)}\big(\mathbb{E}[z_k(t_i, s)] - \mathbb{E}[\mathbb{1}_{\Omega\setminus A} z_k(t_i, s)]\big).$$

$\mathbb{E}[z_k(t_i, s)]$ is affine linear by construction and $\mathbb{E}[\mathbb{1}_{\Omega\setminus A} z_k(t_i, s)]$ is independent of s. Therefore, $\bar{z}(t_i, \cdot)$ is *pathwise affine linear* on the whole interval (s_{j-1}, s_{j+1}). We obtain that

$$\sum_{j=0}^{i-1} x(t_j, s)^\top f_j(x(t_0, s), \dots, x(t_j, s)) + \bar{h}_i(s), \tag{1.15}$$

where

$$\begin{aligned}\bar{h}_i(s) := x(t_i, s)^\top f_i(x(t_0, s), \dots, x(t_i, s)) + \alpha X(t_i, s)^\top \Sigma(t_{i+1}) X(t_i, s) \\ + (X(t_i, s) - x(t_i, s))^\top \mathbb{E}[\tilde{P}(t_i) - \tilde{P}(t_{i+1})] - \mathbb{E}[\bar{z}(t_i, s)^\top (\tilde{P}(t_i) - \tilde{P}(t_{i+1}))] \\ + \mathbb{E}[v(t_{i+1}, x(t_0, s), \dots, x(t_i, s), X(t_i, s) - x(t_i, s) - \bar{z}(t_i, s))]\end{aligned}$$

is strictly convex in s on (s_{j-1}, s_{j+1}) as before by the strict convexity of H_{i+1}. It is clear that $\bar{h}_i(s) = h_i(s)$ for $s \le s_j$. If

$$\bar{h}_i(s) \le h_i(s) \quad \text{for all } s \in (s_{j-1}, s_{j+1}),$$

convexity of $\tilde{h}_i$ follows at the point s_j.

To this end, we observe that for $s > s_j$,

$$\begin{aligned}h_i(s) - \bar{h}_i(s) &= \mathbb{E}[\mathbb{1}_A(z_k(t_i, s) - \bar{z}_k(t_i, s))\epsilon_k(t_{i+1})] \\ &\quad + \mathbb{E}[\mathbb{1}_A v(t_{i+1}, x(t_0, s), \dots, x(t_i, s), X(t_i, s) - x(t_i, s) - z(t_i, s))] \\ &\quad - \mathbb{E}[\mathbb{1}_A v(t_{i+1}, x(t_0, s), \dots, x(t_i, s), X(t_i, s) - x(t_i, s) - \bar{z}(t_i, s))] \\ &=: C_1 + C_2 - C_3.\end{aligned}$$

By Assumption 1.2.1 (ii), $z_1(t_i,s),\dots,z_{k-1}(t_i,s),z_{k+1}(t_i,s),\dots,z_n(t_i,s)$ are independent of A, and thus (using Assumption 1.2.1 (ii) again) $z_1(t_i,s),\dots,z_n(t_i,s)$ are independent with respect to the probability distribution $\mathbb{P}[\cdot|A]$. We denote the distribution of z and z_k with respect to $\mathbb{P}[\cdot|A]$ by $\mathbb{P}^{z|A}$ respectively $\mathbb{P}^{z_k|A}$. By induction hypothesis,

$$v(t_{i+1},x(t_0,s),\dots,x(t_i,s),\cdot)$$

is strictly convex, and thus by Jensen's inequality, the definition of $\bar{z}_k$ on A and the independence of $z_1(t_i,s),\dots,z_n(t_i,s)$ with respect to $\mathbb{P}[\cdot|A]$, we have

$$\begin{aligned}
C_2 &= \mathbb{E}[\mathbb{1}_A v(t_{i+1},x(t_0,s),\dots,x(t_i,s),X(t_i,s)-x(t_i,s)-z(t_i,s))] \\
&= \mathbb{P}[A]\mathbb{E}[v(t_{i+1},x(t_0,s),\dots,x(t_i,s),X(t_i,s)-x(t_i,s)-z(t_i,s))|A] \\
&= \mathbb{P}[A]\int v(t_{i+1},x(t_0,s),\dots,x(t_i,s),X(t_i,s)-x(t_i,s)-a)\mathbb{P}^{z|A}(da) \\
&= \mathbb{P}[A]\int\cdots\int v(t_{i+1},x(t_0,s),\dots,x(t_i,s),X_1(t_i,s)-x_1(t_i,s)-a_1, \\
&\qquad \dots,X_n(t_i,s)-x_n(t_i,s)-a_n)\mathbb{P}^{z_1|A}(da_1)\cdots\mathbb{P}^{z_n|A}(da_n) \\
&= \mathbb{P}[A]\int\cdots\int \mathbb{E}\Big[v(t_{i+1},x(t_0,s),\dots,x(t_i,s),(X_1(t_i,s)-x_1(t_i,s)-a_1, \\
&\qquad \dots,X_k(t_i,s)-x_k(t_i,s)-z_k(t_i,s),\dots,X_n(t_i,s)-x_n(t_i,s)-a_n))|A\Big] \\
&\qquad \mathbb{P}^{z_1|A}(da_1)\cdots\mathbb{P}^{z_{k-1}|A}(da_{k-1})\mathbb{P}^{z_{k+1}|A}(da_{k+1})\cdots\mathbb{P}^{z_n|A}(da_n) \\
&\geq \mathbb{P}[A]\int\cdots\int v(t_{i+1},x(t_0,s),\dots,x(t_i,s),(X_1(t_i,s)-x_1(t_i,s)-a_1, \\
&\qquad \dots,X_k(t_i,s)-x_k(t_i,s)-\mathbb{E}[z_k(t_i,s)|A],\dots,X_n(t_i,s)-x_n(t_i,s)-a_n)) \\
&\qquad \mathbb{P}^{z_1|A}(da_1)\cdots\mathbb{P}^{z_{k-1}|A}(da_{k-1})\mathbb{P}^{z_{k+1}|A}(da_{k+1})\cdots\mathbb{P}^{z_n|A}(da_n) \\
&= \mathbb{P}[A]\mathbb{E}[v(t_{i+1},x(t_0,s),\dots,x(t_i,s),X(t_i,s)-x(t_i,s)-\bar{z}(t_i,s))|A] \\
&= \mathbb{E}[\mathbb{1}_A v(t_{i+1},x(t_0,s),\dots,x(t_i,s),X(t_i,s)-x(t_i,s)-\bar{z}(t_i,s))] \\
&= C_3.
\end{aligned}$$

Note that the above argument fails if the additional Assumption 1.2.1 (ii) is not satisfied.

The last step is to show that $C_1 \geq 0$. We let m_1 be the slope of $y_k(t_i,s)$ for $s<s_j$ and m_2 be the slope of $y_k(t_i,s)$ for $s>s_j$. Note that $0<m_1<m_2$. For $s\leq s_j$, we have $y_k(t_i,s)=\bar{z}_k(t_i,s)$ on A. Thus as $\bar{z}_k(t_i,\cdot)$ is affine linear, $\bar{z}_k(t_i,\cdot)$ has slope m_1 on the whole interval (s_{j-1},s_{j+1}) (cf. Figure 1.2).

Let now $s>s_j$. For

$$A_1 := \{y_k(t_i,s_j)=b(t_i)\}, \quad A_2 := \{y_k(t_i,s_j)<b(t_i)\},$$

we have on $A=A_1\dot{\cup}A_2$,

$$\bar{z}_k(t_i,s)>z_k(t_i,s) \text{ on } A_1, \quad \bar{z}_k(t_i,s)<z_k(t_i,s) \text{ on } A_2.$$

More precisely,

$$z_k(t_i, s) - \bar{z}_k(t_i, s) = -m_1(s - s_j) \quad \text{on } A_1, \tag{1.16}$$
$$z_k(t_i, s) - \bar{z}_k(t_i, s) = (m_2 - m_1)(s - s_j) \quad \text{on } A_2. \tag{1.17}$$

Note that

$$\begin{aligned}
&y_k(t_i, s_j) + m_1(s - s_j) \\
&\quad = \mathbb{E}[z_k(t_i, s)|A] \\
&\quad = \frac{1}{\mathbb{P}[A]}(\mathbb{E}[\mathbb{1}_{A_1} z_k(t_i, s)] + \mathbb{E}[\mathbb{1}_{A_2} z_k(t_i, s)]) \\
&\quad = \frac{1}{\mathbb{P}[A]}(\mathbb{P}[A_1] y_k(t_i, s_j) + \mathbb{P}[A_2] y_k(t_i, s)) \\
&\quad = \frac{1}{\mathbb{P}[A]}((\mathbb{P}[A] - \mathbb{P}[A_2]) y_k(t_i, s_j) + \mathbb{P}[A_2](y_k(t_i, s_j) + m_2(s - s_j))).
\end{aligned} \tag{1.18}$$

Equation (1.18) is equivalent to

$$\begin{aligned}
\mathbb{P}[A_1] &= \frac{m_2 - m_1}{m_2} \mathbb{P}[A], \\
\mathbb{P}[A_2] &= \frac{m_1}{m_2} \mathbb{P}[A].
\end{aligned}$$

Therefore, we have

$$\begin{aligned}
C_1 &= \mathbb{E}\big[\mathbb{1}_A (z_k(t_i, s) - \bar{z}_k(t_i, s)) \epsilon_k(t_{i+1})\big] \\
&= \mathbb{E}\big[\mathbb{1}_{A_1} (z_k(t_i, s) - \bar{z}_k(t_i, s)) \epsilon_k(t_{i+1})\big] + \mathbb{E}\big[\mathbb{1}_{A_2} (z_k(t_i, s) - \bar{z}_k(t_i, s)) \epsilon_k(t_{i+1})\big] \\
&\overset{(1.17),(1.16)}{=} -\mathbb{P}[A_1] m_1 (s - s_j) \mathbb{E}\big[\epsilon_k(t_{i+1})|A_1\big] \\
&\quad + \mathbb{P}[A_2](m_2 - m_1)(s - s_j) \mathbb{E}\big[\epsilon_k(t_{i+1})|A_2\big] \\
&= -\frac{m_1(m_2 - m_1)}{m_2} \mathbb{P}[A](s - s_j) \mathbb{E}\big[\epsilon_k(t_{i+1})|b_k(t_i) = y_k(t_i, s_j)\big] \\
&\quad + \frac{m_1(m_2 - m_1)}{m_2} \mathbb{P}[A](s - s_j) \mathbb{E}\big[\epsilon_k(t_{i+1})|b_k(t_i) > y_k(t_i, s_j)\big] \\
&\geq 0,
\end{aligned} \tag{1.19}$$

where Inequality (1.19) follows from Assumption 1.1.3 (iii).

The cases

$$y(t_i, s_j) < 0 \quad \text{and} \quad y(t_i, s_j) = 0$$

follow similarly with straightforward modifications. Combining these observations, we have strict convexity of $\tilde{h}_i$ at all points s_j and on all intervals (s_j, s_{j+1}), i.e., on the whole interval $[0, 1]$, completing the proof of (iii).

□

1.3. Linear price impact

In the previous section we established an existence and uniqueness result for a general market model. In order to obtain additional insight into the structure of the optimal liquidation strategy we now consider the case of linear temporary price impact, which can be solved in explicit form. In Section 1.3.1 we specify the model in terms of its price impact functions f_i, the fundamental price process $\tilde{P}$ and the liquidity in the dark pool $a(t_0), b(t_0), ..., a(t_N), b(t_N)$. In Section 1.3.2 the value function v and the optimal orders $x^*(t_i), y^*(t_i)$ at times t_i are proven to be of quadratic respectively linear form and shown to satisfy a backward recursion. In Sections 1.3.3 and 1.3.4 we study the effects of dark pools for liquidation of a *single asset position* and a *two asset portfolio*, respectively.

1.3.1. Model description

We do not require Assumption 1.2.1 in this section. The probability space Ω can be infinite, and we allow for dependencies between the dark pool liquidities of the n assets. We therefore do not make use of the findings of Section 1.2.

We specify the impact costs, the distributions of the fundamental asset price $\tilde{P}$ and the dark pool liquidity (a, b) in the following way.

Assumption 1.3.1. *(i) For $i = 0, \dots, N$,*

$$f_i(x(t_i)) := f_i(x(t_0), \dots, x(t_i)) := \Lambda x(t_i)$$

for a positive definite matrix $\Lambda \in \mathbb{R}^{n \times n}$.

(ii) $\tilde{P}$ is an $\mathbb{F}$-martingale and

$$\Sigma(t_i) = \Sigma(t_j) =: \Sigma \quad \textit{for all } i, j = 0, \dots, N.$$

(iii) For $i = 0, \dots, N$, $k = 1, \dots, n$,

$$a_k(t_i), b_k(t_i) \in \{0, \infty\},$$

$\left((a(t_i), b(t_i))\right)_{i=0,\dots,N}$ is identically distributed and $(a(t_i), b(t_i))$ is independent of $\epsilon(t_{i+1})$.

(iv) For $i = 0, \dots, N$, $k = 1, \dots, n$,

$$\begin{aligned}
&\mathbb{P}\big[a_k(t_i) = \infty | a_1(t_i), \dots, a_{k-1}(t_i), a_{k+1}(t_i), \dots, a_n(t_i),\\
&\qquad\qquad b_1(t_i), \dots, b_{k-1}(t_i), b_{k+1}(t_i), \dots, b_n(t_i)\big]\\
&= \mathbb{P}\big[b_k(t_i) = \infty | a_1(t_i), \dots, a_{k-1}(t_i), a_{k+1}(t_i), \dots, a_n(t_i),\\
&\qquad\qquad b_1(t_i), \dots, b_{k-1}(t_i), b_{k+1}(t_i), \dots, b_n(t_i)\big].
\end{aligned}$$

Assumption 1.3.1 (i) implies convexity and superlinear growth of the price impact costs, so that Assumption 1.1.1 is satisfied. We call the matrix Λ the *price impact* *matrix* and say that the price impact is *linear* and *temporary* since the function f_i only depends on the trade $x(t_i)$ at time t_i and not on past trades $x(t_0),...,x(t_{i-1})$. As a direct consequence of Assumption 1.3.1 (i), $v(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i))$ is independent of $x(t_0), \dots, x(t_{i-1})$. The martingale property (Assumption 1.3.1 (ii)), the independence of future price moves of dark pool liquidity (Assumption 1.3.1 (iii)) and the liquidation constraint (Definition 1.1.4 (i)) imply

$$\mathbb{E}\Big[\sum_{j=i}^{N}(x(t_j)+z(t_j))^\top\big(\tilde{P}(t_i)-\tilde{P}(t_j)\big)\Big]=0.$$

Combining the above observations, the Optimization Problem ($\mathrm{OPT}_{\mathrm{dis}}^{\mathrm{gen}}$) becomes

$$\begin{aligned} v(t_i, X(t_i)) &:= v(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i)) \\ &= \inf_{(x,y)\in\mathbb{A}(t_i,X(t_i))}\Big\{\mathbb{E}\Big[\sum_{j=i}^{N}x(t_j)^\top\Lambda x(t_j)\Big]+\alpha\mathbb{E}\Big[\sum_{j=i}^{N}X(t_j)^\top\Sigma X(t_j)\Big]\Big\}. \end{aligned} \tag{$\mathrm{OPT}_{\mathrm{dis}}$}$$

The first part of Assumption 1.3.1 (iii) means that dark pool orders are executed fully or not at all (e.g., in the sense of so-called "fill or kill" orders). The symmetry condition (iv) for dark pool liquidity is required for tractability of the model as we demonstrate in the following remark.

Remark 1.3.2. *Let us illustrate Assumption 1.3.1 (iv) for $n=1$. Assumption 1.3.1 (iii) and (iv) imply for $i=0,\dots,N$,*

$$p := \mathbb{P}\big[a(t_i)=\infty\big]=\mathbb{P}\big[b(t_i)=\infty\big] \tag{1.20}$$

for some $p\in[0,1)$ independent of i. In order to compute the value function of the Optimization Problem ($\mathrm{OPT}_{\mathrm{dis}}$)*, we can proceed by backward induction using the corresponding Bellman equation (cf. the proof of Theorem 1.3.4 below)*

$$\begin{aligned} v(t_i,X(t_i)) &= \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}}\Big\{\Lambda x^2+\alpha\Sigma X(t_i)^2+\mathbb{E}\big[v(t_{i+1},X(t_i)-x-z)\big]\Big\} \\ &\overset{(1.20)}{=} \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}}\Big\{\Lambda x^2+\alpha\Sigma X(t_i)^2+pv(t_{i+1},X(t_i)-x-y) \\ &\qquad + (1-p)v(t_{i+1},X(t_i)-x)\Big\} \\ &=: \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}}\tilde{v}(x,y). \end{aligned} \tag{1.21}$$

Strict convexity of $v(t_i,\cdot)$ now carries over to strict convexity of $\tilde{v}$. If $v(t_{i+1},\cdot)$ is known explicitly (and of sufficiently simple form), computing $v(t_i,X(t_i))$ becomes a standard minimization problem.

If we drop Assumption 1.3.1 (iv), the situation becomes uncomfortable. We obtain

$$p_1 := \mathbb{P}\big[a(t_i)=\infty, b(t_i)=\infty\big], \quad p_2 := \mathbb{P}\big[a(t_i)=\infty, b(t_i)=0\big],$$

$$p_3 := \mathbb{P}\big[a(t_i) = 0, b(t_i) = \infty\big], \quad p_4 := \mathbb{P}\big[a(t_i) = 0, b(t_i) = 0\big] = 1 - p_1 - p_2 - p_3$$

for some $p_1, p_2, p_3, p_4 \in [0, 1)$ *instead of Equation* (1.20) *and thus*

$$\begin{aligned} v(t_i, X(t_i)) = \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}} \Big\{ & \Lambda x^2 + \alpha\Sigma X(t_i)^2 \\ & + \mathbb{1}_{\{y\geq 0\}}\big((p_1+p_3)v(t_{i+1}, X(t_i) - x - y) + (p_2+p_4)v(t_{i+1}, X(t_i) - x)\big) \\ & + \mathbb{1}_{\{y<0\}}\big((p_1+p_2)v(t_{i+1}, X(t_i) - x - y) + (p_3+p_4)v(t_{i+1}, X(t_i) - x)\big)\Big\} \\ =: \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}} & \tilde{v}(x, y) \end{aligned}$$

instead of Equation (1.21)*. The resulting optimization problem is thus significantly more complicated. Additionally, it can be shown that strict convexity of* $v(t_{i+1}, \cdot)$ *does not transfer to strict convexity of* $\tilde{v}$ *for* $n \geq 2$ *in general. We omit the execution here.*

Before we solve the Optimization Problem ($\mathrm{OPT}_{\mathrm{dis}}$), we need to introduce the following notation. In each time-interval $[t_i, t_{i+1})$ there are 2^n possible "combinations" or "scenarios" with respect to joint execution and non-execution of the order $y(t_i) \in \mathbb{R}^n$ in the dark pool (note that due to Assumption 1.3.1 (iv), we can assume that execution of an order $y_k(t_i)$ is "independent" of the sign of $y_k(t_i)$). Each of these scenarios occurs with a fixed probability, which we denote by p_l for the l^{th} scenario, determined by the distributions of the random variables $a(t_i)$ and $b(t_i)$. We denote the amount executed in the dark pool at time i in scenario l by $z(l, t_i)$, i.e.,

$$z_k(l, t_i) := \begin{cases} y_k(t_i) & \text{if in the } l^{\text{th}} \text{ scenario the order in the } k^{\text{th}} \text{ asset} \\ & \text{in the dark pool is executed} \\ 0 & \text{otherwise.} \end{cases}$$

There exists a diagonal matrix $Z_l \in \mathbb{R}^{n\times n}$ (with 1's and 0's on the diagonal) such that

$$z(l, t_i) = Z_l y(t_i).$$

We define the diagonal matrix $\hat{P} = (\hat{p}_{k,m})_{k,m=1,\ldots,n}$ by

$$\hat{P} := \sum_l p_l Z_l, \tag{1.22}$$

i.e., $\hat{p}_{k,k}$ is the probability that an order for the k^{th} asset is executed in the dark pool in $[t_i, t_{i+1})$. We re-order the assets in such a way that for $k_0 \in \{0, \ldots, n\}$,

$$\hat{p}_{k,k} = 0 \quad \text{if and only if} \quad k > k_0,$$

i.e., $k_0 = 0$ refers to the case where the dark pool is not used at all and $k_0 = n$ to the case where there is liquidity with positive probability for all assets in the dark pool.

Finally, we introduce the following notation.

Notation 1.3.3. *For a positive definite matrix $M \in \mathbb{R}^{n\times n}$, we define*

$$\check{M} = (\check{m}_{i,j})_{i,j=1,\dots,n} := \sum_l p_l Z_l M Z_l.$$

For independent liquidity of the n assets in the dark pool, i.e., if Assumption 1.2.1 (ii) holds, $\check{M}$ becomes significantly simpler:

$$\check{m}_{i,j} = \begin{cases} m_{i,j}\hat{p}_{i,i} & \text{if } i = j \\ m_{i,j}\hat{p}_{i,i}\hat{p}_{j,j} & \text{if } i \neq j. \end{cases}$$

Note that both $\check{M} = (\check{m}_{k,m})_{k,m=1,\dots,n}$ and $\hat{P}$ are positive definite for $k_0 = n$ but not for $k_0 < n$. In the latter case, $\check{m}_{k,m} = 0$ for $k > k_0$ or $m > k_0$ and $\hat{p}_{k,k} = 0$ for $k > k_0$. However, the matrices $\check{M}' := (\check{m}_{k,m})_{k,m=1,\dots,k_0} \in \mathbb{R}^{k_0\times k_0}$ and $\hat{P}' := (\hat{p}_{k,m})_{k,m=1,\dots,k_0} \in \mathbb{R}^{k_0\times k_0}$ are positive definite. We therefore use generalized inverses of matrices. We denote the Moore-Penrose Inverse of a matrix M by $M^\dagger$. For regular M, we have $M^{-1} = M^\dagger$ (see, e.g., the book by Ben-Israel and Greville [2003]).

1.3.2. Optimal liquidation

We are now able to solve the Optimization Problem ($\mathrm{OPT}_{\mathrm{dis}}$). It turns out that the value function $v(t_i, \cdot)$ is quadratic and that the optimal orders $x^*(t_i)$ placed at the primary venue and $y^*(t_i)$ placed in the dark pool are linear functions of the portfolio $X(t_i)$ at any time t_i.

Theorem 1.3.4. *Let $i = 0, \dots, N$ and assume that $X(t_i) \in \mathbb{R}^n$ is the portfolio position at time t_i. Then there exists a unique optimal strategy $(x^*, y^*) \in \mathbb{A}(t_i, X(t_i))$ realizing the minimum in Equation* ($\mathrm{OPT}_{\mathrm{dis}}$).

Moreover, there exist matrices $A(t_i), B(t_i), C(t_i) \in \mathbb{R}^{n\times n}$ independent of $X(t_i)$ such that the optimal strategy fulfills

$$x^*(t_i) = A(t_i)X(t_i),$$
$$y^*(t_i) = B(t_i)X(t_i),$$

and the value function is given by

$$v(t_i, X(t_i)) = X(t_i)^\top C(t_i) X(t_i)$$

with positive definite $C(t_i)$. $A(t_i), B(t_i)$ and $C(t_i)$ are given recursively by $A(t_N) = I$, $B(t_N) = 0$, $C(t_N) = \Lambda + \alpha\Sigma$ and

$$A(t_i) = \big(\Lambda + D(t_{i+1})\big)^\dagger D(t_{i+1}), \tag{1.23}$$

$$B(t_i) = \check{C}(t_{i+1})^\dagger \hat{P} C(t_{i+1})\big(I - A(t_i)\big), \tag{1.24}$$

$$C(t_i) = \alpha\Sigma + D(t_{i+1}) - D(t_{i+1})\big(\Lambda + D(t_{i+1})\big)^\dagger D(t_{i+1}), \tag{1.25}$$

where

$$D(t_{i+1}) := C(t_{i+1}) - C(t_{i+1})\hat{P}\check{C}(t_{i+1})^{\dagger}\hat{P}C(t_{i+1}). \tag{1.26}$$

Proof. We prove the theorem by backward induction. Note that $C(t_N) = \Lambda + \alpha\Sigma$ is positive definite as Λ is positive definite, Σ (as a covariance matrix) is nonnegative definite and $\alpha \geq 0$. Therefore all assertions follow for $i = N$.

Let now $i < N$. Due to the linearity of the price impact function, the martingale property and the independence of dark pool liquidity of future price moves, we obtain the Bellman equation

$$\begin{aligned} v(t_i, X(t_i)) = \inf_{\substack{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n \\ y_k=0,\, k>k_0}} \Big\{ & x^\top \Lambda x + \alpha \cdot X(t_i)^\top \Sigma X(t_i) \\ & + \sum_l p_l v(t_{i+1}, X(t_i) - x - Z_l y) \Big\}; \end{aligned} \tag{1.27}$$

recall that the optimal strategy must fulfill $y_k(t_i) = 0$ for $k > k_0$ almost surely by Assumption 1.1.4 (ii). By abuse of notation, y denotes simultaneously $y \in \mathbb{R}^{k_0}$ and $y \in \mathbb{R}^n$ with $y_k = 0$ for $k > k_0$, where either way is clear from the context. Using the induction hypothesis, Equation (1.27) becomes

$$\begin{aligned} & v(t_i, X(t_i)) \\ & = \inf_{\substack{(x,y)\in\mathbb{R}^n\times\mathbb{R}^n \\ y_k=0,\, k>k_0}} \Big\{ x^\top \Lambda x + \alpha X(t_i)^\top \Sigma X(t_i) \\ & \qquad + \sum_l p_l (X(t_i) - x - Z_l y)^\top C(t_{i+1})(X(t_i) - x - Z_l y) \Big\}. \end{aligned}$$

The function $\tilde{v}(t_i, X(t_i), \cdot) : \mathbb{R}^n \times \mathbb{R}^{k_0} \longrightarrow \mathbb{R}$ given by

$$\begin{aligned} \tilde{v}(t_i, X(t_i), x, y) := {} & x^\top \Lambda x + \alpha \cdot X(t_i)^\top \Sigma X(t_i) \\ & + \sum_l p_l (X(t_i) - x - Z_l y)^\top C(t_{i+1})(X(t_i) - x - Z_l y) \end{aligned}$$

is a strictly convex linear-quadratic functional as $C(t_{i+1})$ is positive definite by the induction hypothesis and Λ is positive definite. Therefore, the unique minimum (x^*, y^*) of $\tilde{v}(t_i, X(t_i), \cdot)$ is given by the solution of

$$\begin{aligned} \nabla_x \tilde{v}(t_i, X(t_i), x, y) &= 0, \\ \nabla_y \tilde{v}(t_i, X(t_i), x, y) &= 0. \end{aligned} \tag{1.28}$$

The System (1.28) is equivalent to

$$\begin{aligned} (\Lambda + C(t_{i+1}))x + C(t_{i+1})\hat{P}y &= C(t_{i+1})X(t_i), \\ \hat{P}C(t_{i+1})x + \check{C}(t_{i+1})y &= \hat{P}C(t_{i+1})X(t_i), \end{aligned}$$

which in turn is equivalent to

$$
\begin{aligned}
\hat{P}C(t_{i+1})x + \check{C}(t_{i+1})y &= \hat{P}C(t_{i+1})X(t_i), \\
(\Lambda + D(t_{i+1}))x &= D(t_{i+1})X(t_i).
\end{aligned}
\tag{1.29}
$$

Solving System (1.29) for (x, y), yields Equations (1.23) and (1.24). Plugging this into $\tilde{v}(t_i, X(t_i), \cdot)$, we obtain

$$
\begin{aligned}
&v(t_i, X(t_i)) \\
&\quad = (A(t_i)X(t_i))^\top \Lambda (A(t_i)X(t_i)) + \alpha \cdot X(t_i)^\top \Sigma X(t_i) \\
&\qquad + \sum_l \Big(p_l (X(t_i) - A(t_i)X(t_i) - Z_l B(t_i)X(t_i))^\top \\
&\qquad\qquad C(t_{i+1})(X(t_i) - A(t_i)X(t_i) - Z_l B(t_i)X(t_i)) \Big) \\
&\quad = X(t_i)^\top \underbrace{\Big(A(t_i)^\top \Lambda A(t_i) + \alpha\Sigma + \sum_l p_l (I - A(t_i) - Z_l B(t_i))^\top C(t_{i+1})(I - A(t_i) - Z_l B(t_i)) \Big)}_{=:C(t_i)} X(t_i).
\end{aligned}
$$

By the induction hypothesis, $C(t_i)$ is nonnegative definite. To see that $C(t_i)$ is indeed positive definite, let $x \in \mathbb{R}^n$, $x_k \neq 0$. If $A(t_i)x \neq 0$, then

$$
x^\top A(t_i)^\top \Lambda A(t_i) x > 0.
$$

In any case, there exists an l such that the k^{th} diagonal element of Z_l is 0 and $p_l > 0$ (cf. Assumption 1.1.3 (ii)). If $A(t_i)x = 0$, then $y := (I - A(t_i) - Z_l B(t_i))x \neq 0$ and

$$
p_l x^\top (I - A(t_i) - Z_l B(t_i))^\top C(t_{i+1})(I - A(t_i) - Z_l B(t_i))x = p_l y^\top C(t_{i+1}) y > 0
$$

by the induction hypothesis. We now set

$$
E(t_{i+1}) := C(t_{i+1})\hat{P}\check{C}(t_{i+1})^\dagger \hat{P}C(t_{i+1}) = C(t_{i+1}) - D(t_{i+1}).
$$

Note that $\check{C}(t_{i+1})$ and thus $D(t_{i+1})$ and $E(t_{i+1})$ are symmetric due to the symmetry of $C(t_{i+1})$. Furthermore,

$$
\begin{aligned}
C(t_i) &= A(t_i)^\top \Lambda A(t_i) + \alpha\Sigma + C(t_{i+1}) - C(t_{i+1})A(t_i) - A(t_i)^\top C(t_{i+1}) \\
&\quad - C(t_{i+1})\hat{P}B(t_i) - B(t_i)^\top \hat{P}C(t_{i+1}) + A(t_i)^\top C(t_{i+1})A(t_i) \\
&\quad + B(t_i)^\top \check{C}_{i+1} B(t_i) + A(t_i)^\top C(t_{i+1})\hat{P}B(t_i) + B(t_i)^\top \hat{P}C(t_{i+1})A(t_i) \\
&= A(t_i)^\top \Lambda A(t_i) + \alpha\Sigma + C(t_{i+1}) - C(t_{i+1})A(t_i) - A(t_i)^\top C(t_{i+1}) - E(t_{i+1}) \\
&\quad + A(t_i)^\top C(t_{i+1})A(t_i) + E(t_{i+1})A(t_i) + A(t_i)^\top E(t_{i+1}) - A(t_i)^\top E(t_{i+1})A(t_i)
\end{aligned}
$$

$$
\begin{aligned}
&= A(t_i)^\top \Lambda A(t_i) + \alpha\Sigma + D(t_{i+1}) - D(t_{i+1})A(t_i) - A(t_i)^\top D(t_{i+1}) \\
&\quad + A(t_i)^\top D(t_{i+1})A(t_i) \\
&= \alpha\Sigma + D(t_{i+1}) - D(t_{i+1})(\Lambda + D(t_{i+1}))^\dagger D(t_{i+1})
\end{aligned}
$$

as required. □

1.3.3. Liquidating a single asset position

The most transparent case to analyze is the liquidation of a position $X(t_0)$ in a *single asset* ($n = 1$), for which we derive a closed form solution. Some of the most interesting effects of using dark pools can be observed in this case, and we therefore study it in more depth in this section.

Let

$$p := \mathbb{P}[a(t_i) = \infty] = \mathbb{P}[b(t_i) = \infty]$$

be the probability of order execution in the dark pool. For a given $p \in [0, 1)$, we denote the matrices (now real numbers) $A(t_i), B(t_i)$ and $C(t_i)$ introduced in Theorem 1.3.4 by $A(t_i, p), B(t_i, p)$ and $C(t_i, p)$ in order to highlight their dependence on p.

For the remainder of the section, we let

$$X(t_0) > 0.$$

The results are symmetric in the sign of $X(t_0)$ and can be easily transfered to negative initial asset positions $X(t_0)$.

Closed form solutions for the value function and the optimal strategy

The following proposition gives closed form solutions for the optimal strategy and the value function. This result is a generalization of the corresponding result in Almgren and Chriss [2001] for optimal liquidation without dark pools ($p = 0$).

Proposition 1.3.5. *Let*

$$\kappa(p) := \operatorname{arcosh}\Big(\frac{\sqrt{1-p}}{2}\big(\frac{\alpha\Sigma}{\Lambda} + 1 + \frac{1}{1-p}\big)\Big). \tag{1.30}$$

Then the optimal orders at time t_i are given by

$$x^*(t_i) = A(t_i, p)X(t_i) \quad \text{and} \quad y^*(t_i) = B(t_i, p)X(t_i)$$

with

$$A(t_i, p) = 1 - \frac{\sinh(\kappa(p)(N-i))}{\sqrt{1-p}\,\sinh(\kappa(p)(N+1-i))}, \tag{1.31}$$

$$B(t_i, p) = 1 - A(t_i, p) = \frac{\sinh(\kappa(p)(N-i))}{\sqrt{1-p}\,\sinh(\kappa(p)(N+1-i))} < 1. \tag{1.32}$$

In particular,

$$0 < x^*(t_i), y^*(t_i) < X(t_i) \quad \textit{and} \quad x^*(t_i) + y^*(t_i) = X(t_i)$$

for $i \neq N$*. The value function is given by*

$$v(t_i, X(t_i)) = C(t_i, p) X(t_i)^2$$

with

$$C(t_i, p) = \frac{\Lambda}{1-p}\Big(\frac{\sqrt{1-p}\sinh(\kappa(p)(N+2-i))}{\sinh(\kappa(p)(N+1-i))} - 1\Big). \tag{1.33}$$

Proof. Note first that for $i < N$, we have $\check{C}(t_{i+1}, p) = pC(t_{i+1}, p)$ and $D(t_{i+1}, p) = (1-p)C(t_{i+1}, p)$. Thus by Equations (1.23), (1.24) and (1.25),

$$A(t_i, p) = \frac{(1-p)C(t_{i+1}, p)}{(1-p)C(t_{i+1}, p) + \Lambda}, \quad B(t_i, p) = \frac{\Lambda}{(1-p)C(t_{i+1}, p) + \Lambda} \tag{1.34}$$

and

$$C(t_i, p) = \frac{\alpha\Sigma\Lambda + (1-p)C(t_{i+1}, p)(\Lambda + \alpha\Sigma)}{\Lambda + (1-p)C(t_{i+1}, p)}. \tag{1.35}$$

We therefore define

$$u_k := C(t_{N-k}, p) > 0, \quad k = 0, \dots, N$$

and obtain the following recursion:

$$\begin{aligned} u_0 &= \Lambda + \alpha\Sigma, \\ u_{k+1} &= \frac{\alpha\Sigma\Lambda + (1-p)u_k(\Lambda + \alpha\Sigma)}{\Lambda + (1-p)u_k}. \end{aligned}$$

This recursion can be solved as follows. We set

$$w_k := \frac{\sqrt{1-p}}{\Lambda}\Big(u_k + \frac{\Lambda}{1-p}\Big).$$

Since $\kappa(p)$ is well-defined and strictly positive as

$$\frac{\sqrt{1-p}}{2}\Big(\frac{\alpha\Sigma}{\Lambda} + 1 + \frac{1}{1-p}\Big) \geq \frac{1}{2}\Big(\sqrt{1-p} + \frac{1}{\sqrt{1-p}}\Big) > 1,$$

we obtain the following modified recursion:

$$\begin{aligned} w_0 &= \frac{\sqrt{1-p}}{\Lambda}\Big(\Lambda + \alpha\Sigma + \frac{\Lambda}{1-p}\Big) \\ &= 2\cosh(\kappa(p)), \\ w_{k+1} &= \frac{\sqrt{1-p}}{\Lambda}\Big(\frac{\Lambda}{1-p} + u_{k+1}\Big) \end{aligned}$$

$$
\begin{aligned}
&= \frac{\sqrt{1-p}}{\Lambda}\Big(\frac{\Lambda}{1-p} + \frac{\alpha\Sigma\Lambda + (1-p)(\Lambda+\alpha\Sigma)\big(\frac{\Lambda}{\sqrt{1-p}}w_k - \frac{\Lambda}{1-p}\big)}{\Lambda + (1-p)\big(\frac{\Lambda}{\sqrt{1-p}}w_k - \frac{\Lambda}{1-p}\big)}\Big)\\
&= 2\cosh(\kappa(p)) - \frac{1}{w_k}.
\end{aligned}
$$

This well-known type of recursion (see, e.g., Perron [1954], §12) can be solved explicitly: if

$$a := \kappa(p), \quad b := (k+2)\kappa(p)$$

and

$$w_k = \frac{\sinh\left((k+2)\kappa(p)\right)}{\sinh\left((k+1)\kappa(p)\right)},$$

then

$$w_0 = \frac{\sinh\left(2\kappa(p)\right)}{\sinh\left(\kappa(p)\right)} = 2\cosh\left(\kappa(p)\right)$$

and

$$
\begin{aligned}
w_{k+1} + \frac{1}{w_k} =& \frac{\sinh(b+a)}{\sinh(b)} + \frac{\sinh(b-a)}{\sinh(b)}\\
=& \frac{\exp(b+a) - \exp(-b-a) + \exp(b-a) - \exp(a-b)}{\exp(b) - \exp(-b)}\\
=& \,2\cosh(a)
\end{aligned}
$$

as desired. Equation (1.33) follows directly.

Equations (1.31) and (1.32) follow easily by plugging Equation (1.33) into (1.34). Finally, $B(t_i, p) < 1$ follows directly from Equation (1.34) and the fact that $C(t_{i+1}, p) > 0$. □

We can also express $X(t_i)$, $x^*(t_i)$ and $y^*(t_i)$ as functions of $X(t_0)$.

Corollary 1.3.6. *Let $i = 0, \dots, N$ and assume that no dark pool order has been executed until time t_i. We recursively define*

$$X^{\mathrm{ne}}(t_0, p) := X(t_0), \quad X^{\mathrm{ne}}(t_j, p) := X^{\mathrm{ne}}(t_{j-1}, p) - x^*(t_{j-1}).$$

Then

$$X(t_i) = X^{\mathrm{ne}}(t_i, p) = \frac{1}{\sqrt{1-p}^{\,i}} \frac{\sinh(\kappa(p)(N+1-i))}{\sinh(\kappa(p)(N+1))} X(t_0), \tag{1.36}$$

and the optimal orders at time t_i are given by

$$
\begin{aligned}
x^{\mathrm{ne}}(t_i, p) &:= x^*(t_i)\\
&= A(t_i, p) X(t_i)\\
&= \frac{1}{\sqrt{1-p}^{\,i+1}} \frac{\sqrt{1-p}\sinh(\kappa(p)(N+1-i)) - \sinh(\kappa(p)(N-i))}{\sinh(\kappa(p)(N+1))} X(t_0),
\end{aligned} \tag{1.37}
$$

$$\begin{aligned}
y^{\mathrm{ne}}(t_i,p) &:= y^*(t_i)\\
&= B(t_i,p)X(t_i)\\
&= X^{\mathrm{ne}}(t_i,p) - x^*(t_i,p)\\
&= \frac{1}{\sqrt{1-p}^{i+1}} \frac{\sinh(\kappa(p)(N-i))}{\sinh(\kappa(p)(N+1))} X(t_0).
\end{aligned} \tag{1.38}$$

If a dark pool order has been executed before time t_i, then $X(t_i) = x^*(t_i) = y^*(t_i) = 0$.

Proof. We prove Equation (1.36) by forward induction. Equations (1.37) and (1.38) then follow as

$$\begin{aligned}
x^{\mathrm{ne}}(t_i,p) &= X^{\mathrm{ne}}(t_i,p) - X^{\mathrm{ne}}(t_{i+1},p) \quad \text{and}\\
y^{\mathrm{ne}}(t_i,p) &= X^{\mathrm{ne}}(t_i,p) - x^{\mathrm{ne}}(t_i,p)
\end{aligned}$$

by Proposition 1.3.5.

The case $i = 0$ is clear. For $i > 0$, we use the induction hypothesis and Equation (1.31) and obtain

$$\begin{aligned}
X^{\mathrm{ne}}(t_{i+1},p) &= X^{\mathrm{ne}}(t_i,p) - A(t_i,p)X^{\mathrm{ne}}(t_i,p)\\
&= \frac{\sinh(\kappa(p)(N+1-i))}{\sqrt{1-p}^{i}\sinh(\kappa(p)(N+1))} X(t_0)\\
&\quad - \Big(1 - \frac{\sinh(\kappa(p)(N-i))}{\sqrt{1-p}\sinh(\kappa(p)(N+1-i))}\Big) \frac{\sinh(\kappa(p)(N+1-i))}{\sqrt{1-p}^{i}\sinh(\kappa(p)(N+1))} X(t_0)\\
&= \frac{\sinh(\kappa(p)(N-i))}{\sqrt{1-p}^{i+1}\sinh(\kappa(p)(N+1))} X(t_0)
\end{aligned}$$

as required. □

Properties of the optimal strategy

Proposition 1.3.5 answers the question of how to use a dark pool optimally for $n = 1$. It is always optimal to place the largest possible order in the dark pool as we do not pay price impact there. Consequently, the liquidation task is finished as soon as the dark pool order is executed by Assumption 1.3.1 (iii).

Using the dark pool also changes optimal trading in the primary exchange. Intuitively, the trader should slow down the trading speed in the primary venue as she wants as much as possible to be executed in the dark pool. If the position is not yet executed towards the end, she has to speed up in order to finish the liquidation until time T.

The first part of the following proposition confirms this intuition. The second part states that the expected asset positions for trading with dark pool are smaller than without dark pool, although trading at the primary exchange is slower.

Proposition 1.3.7. *Let $i = 1, ..., N-1$.*

(i) $X^{\text{ne}}(t_i, p)$ *is strictly increasing in* p*. In particular*

$$X^{\text{ne}}(t_i, p) > X^{\text{ne}}(t_i, 0) \quad \textit{for all } p > 0.$$

(ii) Let $\mathbb{E}_p[X(t_i)]$ *be the expected asset position at time* t_i *if the probability of execution in the dark pool is* p *and the optimal strategy is applied:*

$$\mathbb{E}_p[X(t_i)] = (1-p)^i X^{\text{ne}}(t_i, p) = \frac{\sqrt{1-p}^i \sinh(\kappa(p)(N+1-i))}{\sinh(\kappa(p)(N+1))} X(t_0). \tag{1.39}$$

Then $\mathbb{E}_p[X(t_i)]$ *is strictly decreasing in* p*, in particular*

$$\mathbb{E}_p[X(t_i)] < \mathbb{E}_0[X(t_i)] \quad \textit{for all } p > 0.$$

We require the following lemma for the proof of Proposition 1.3.7. Assertion and proof of the lemma are elementary. However, the first part of the lemma plays an important role not only in the proof of Proposition 1.3.7 but also later in the proofs of Proposition 1.3.9, where we prove monotonicity results for the cost components of the optimal strategy and of Proposition 2.5.1, where we prove the corresponding properties for the optimal strategy and the value function in the continuous-time setting of Chapter 2.

Lemma 1.3.8. *(i) Let* $0 < a < b$, $x > 0$*. Then*

$$0 > \frac{d}{dx}\frac{\sinh(ax)}{\sinh(bx)} > (a-b)\frac{\sinh(ax)}{\sinh(bx)}. \tag{1.40}$$

(ii) Let $p \in (0,1)$*. Then*

$$\kappa'(p) = \frac{(1-p)^{-3/2}}{4\Lambda \sinh(\kappa(p))}(p\Lambda - (1-p)\alpha\Sigma), \tag{1.41}$$

in particular

$$\kappa'(p) < 0 \quad \textit{on} \quad (0, \frac{\alpha\Sigma}{\alpha\Sigma + \Lambda}) \quad \textit{and} \quad \kappa'(p) > 0 \quad \textit{on} \quad (\frac{\alpha\Sigma}{\alpha\Sigma+\Lambda}, 1). \tag{1.42}$$

Furthermore,

$$|\kappa'(p)| \le \frac{1}{2(1-p)}. \tag{1.43}$$

Proof. (i) Note first that

$$\frac{d}{dx}\frac{\sinh(ax)}{\sinh(bx)} = \frac{a\cosh(ax)\sinh(bx) - b\cosh(bx)\sinh(ax)}{\sinh^2(bx)} \tag{1.44}$$

$$= \frac{\sinh(ax)}{\sinh(bx)}\Big(a\frac{\cosh(ax)}{\sinh(ax)} - b\frac{\cosh(bx)}{\sinh(bx)}\Big) \tag{1.45}$$

and
$$\frac{d}{dx}\Big(a\frac{\cosh(ax)}{\sinh(ax)} - b\frac{\cosh(bx)}{\sinh(bx)}\Big) = \frac{b^2}{\sinh^2(bx)} - \frac{a^2}{\sinh^2(ax)} < 0$$
as
$$0 < \frac{\sinh(cx)}{c} = \sum_{i=0}^{\infty} \frac{x^{2i+1}c^{2i}}{(2i+1)!}$$
increases strictly in $c > 0$. We obtain
$$\begin{aligned} 0 &= \lim_{t\to 0+}\Big(a\frac{\cosh(at)}{\sinh(at)} - b\frac{\cosh(bt)}{\sinh(bt)}\Big) \\ &> \Big(a\frac{\cosh(ax)}{\sinh(ax)} - b\frac{\cosh(bx)}{\sinh(bx)}\Big) &(1.46)\\ &> \lim_{t\to\infty}\Big(a\frac{\cosh(at)}{\sinh(at)} - b\frac{\cosh(bt)}{\sinh(bt)}\Big) = a - b. &(1.47)\end{aligned}$$

Inequality (1.46) directly implies the first inequality in (1.40) (cf. Equation (1.45)). The second inequality in (1.40) follows from Equation (1.45) and Inequality (1.47).

(ii) We have
$$2\sinh(\kappa(p))\kappa'(p) = 2\frac{d}{dp}\cosh(\kappa(p)) = \frac{(1-p)^{-3/2}}{2} - \frac{(1-p)^{-1/2}}{2}\Big(\frac{\alpha\Sigma}{\Lambda}+1\Big)$$
and thus
$$\kappa'(p) = \frac{(1-p)^{-3/2}}{4\Lambda\sinh(\kappa(p))}\big(p\Lambda - (1-p)\alpha\Sigma\big).$$
Furthermore,
$$\begin{aligned}\sinh^2(\kappa(p)) = \cosh^2(\kappa(p)) - 1 &= \frac{1-p}{4}\Big(\frac{\alpha\Sigma}{\Lambda} + 1 + \frac{1}{1-p}\Big)^2 - 1 \\ &= \frac{1-p}{4}\Big(\underbrace{\big(\frac{1}{1-p}+1\big)^2 - \frac{4}{1-p}}_{=\left(\frac{p}{1-p}\right)^2} + \frac{\alpha^2\Sigma^2}{\Lambda^2} + 2\underbrace{\big(\frac{1}{1-p}+1\big)}_{\geq\frac{p}{1-p}}\frac{\alpha\Sigma}{\Lambda}\Big) \\ &\geq \frac{1-p}{4}\Big(\frac{\alpha\Sigma}{\Lambda} + \frac{1}{1-p} - 1\Big)^2.\end{aligned}$$
Therefore,
$$\sinh(\kappa(p)) \geq \frac{1}{2\Lambda\sqrt{1-p}}\big(p\Lambda + (1-p)\alpha\Sigma\big) \geq \frac{1}{2\Lambda\sqrt{1-p}}\big|p\Lambda - (1-p)\alpha\Sigma\big|, \quad (1.48)$$
hence (cf. Equation (1.41))
$$|\kappa'(p)| = \frac{(1-p)^{-3/2}}{4\Lambda}\frac{|p\Lambda - (1-p)\alpha\Sigma|}{\sinh(\kappa(p))} \leq \frac{1}{2(1-p)}.$$

□

We are now ready to prove Proposition 1.3.7.

Proof of Proposition 1.3.7. Let $0 < a < b$. By Lemma 1.3.8, we obtain

$$\Big|\frac{\partial}{\partial p}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\Big| = \Big|\kappa'(p)\frac{\partial}{\partial \kappa}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\Big| < \frac{b-a}{2(1-p)}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}. \tag{1.49}$$

Therefore,

$$\begin{aligned}
&\frac{\partial}{\partial p}\Big((1-p)^{\frac{b-a}{2}}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\Big)\\
&\quad= -\frac{b-a}{2}(1-p)^{\frac{b-a}{2}-1}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} + (1-p)^{\frac{b-a}{2}}\frac{\partial}{\partial p}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\\
&\quad\overset{(1.49)}{<} -\frac{b-a}{2}(1-p)^{\frac{b-a}{2}-1}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} + \frac{b-a}{2}(1-p)^{\frac{b-a}{2}-1}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\\
&\quad= 0
\end{aligned} \tag{1.50}$$

and similarly

$$\begin{aligned}
&\frac{\partial}{\partial p}\Big((1-p)^{-\frac{b-a}{2}}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\Big)\\
&\quad= \frac{b-a}{2}(1-p)^{-\frac{b-a}{2}-1}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} + (1-p)^{-\frac{b-a}{2}}\frac{\partial}{\partial p}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\\
&\quad\overset{(1.49)}{>} \frac{b-a}{2}(1-p)^{-\frac{b-a}{2}-1}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} - \frac{b-a}{2}(1-p)^{-\frac{b-a}{2}-1}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\\
&\quad= 0.
\end{aligned} \tag{1.51}$$

We can directly deduce the two assertions.

Let $i > 0$ and set $a = N+1-i$ and $b = N+1$, thus yielding $b - a = i$. By Equation (1.36) and Inequality (1.51), we obtain

$$\frac{\partial}{\partial p}X^{\mathrm{ne}}(t_i, p) = X(t_0)\frac{\partial}{\partial p}\Big((1-p)^{-\frac{i}{2}}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\Big) > 0.$$

By Equation (1.39) and Inequality (1.50), we obtain

$$\frac{\partial}{\partial p}\mathbb{E}_p[X(t_i)] = X(t_0)\frac{\partial}{\partial p}\Big((1-p)^{\frac{i}{2}}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\Big) < 0.$$

□

Figures 1.3 and 1.4 illustrate how the dark pool changes the optimal strategy in the primary venue. In all pictures the optimal strategy without using the dark pool is displayed by the thin line. When the dark pool is used, then the portfolio evolution

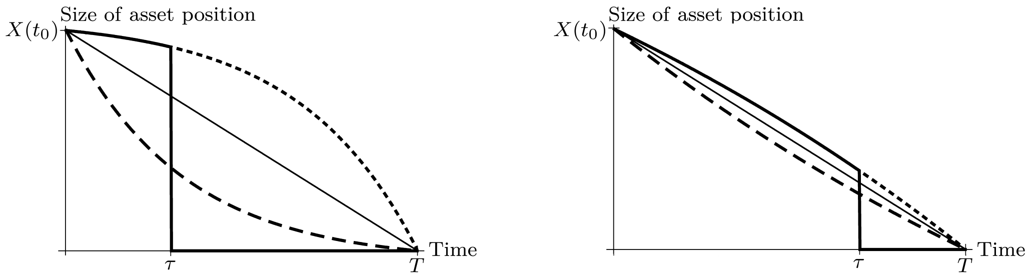


Figure 1.3.: Comparison of the portfolio evolution for the optimal strategies of a risk-neutral trader ($\alpha = 0$) of a stock with large (left) and small (right) probability of execution p in the dark pool. The task of the trader is to liquidate a position $X(t_0) = 1$ in $N+1 = 501$ trading times. Furthermore, $\Lambda = 500$ and $p = \frac{3}{500}$ (left graph), $p = \frac{1}{1000}$ (right graph). The solid thick lines show the scenario where a trade in the dark pool is executed in the τ^{th} trading period (left $\tau = 150$, right $\tau = 350$). Dotted lines refer to the optimal strategy until execution, solid thin lines to the optimal strategy without dark pool and dashed lines to the expected asset position when the dark pool is used.

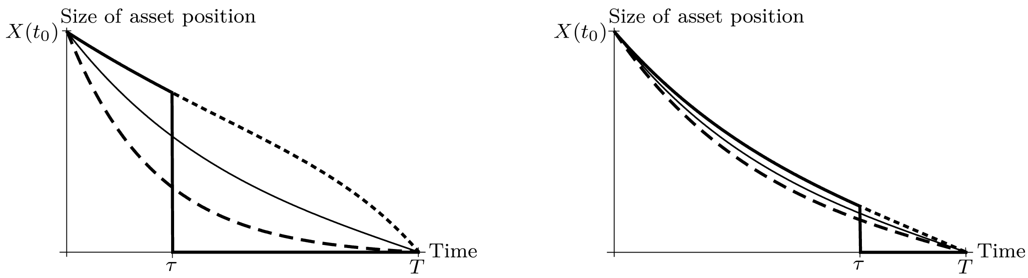


Figure 1.4.: The same liquidation problems as in Figure 1.3 but for a risk-averse trader ($\alpha = 4$, $\Sigma = \frac{1}{500}$).

is stochastic and depends on the liquidity found in the dark pool. We illustrate the stochastic portfolio evolution with three lines. The solid line shows the evolution of the asset position when liquidity is found in the dark pool at time τ. If there is no liquidity found in the dark pool during the entire trading horizon, the trader follows the dotted line until time T. The figures illustrate how the trading speed is slowed down by the introduction of the dark pool. The dashed lines denote the expected asset position over time if the dark pool is used.

Properties of the value function

The costs of an admissible liquidation strategy (x, y) are composed of the *impact costs* of trading at the primary venue

$$\Lambda \cdot \mathbb{E}\Big[\sum_{i=0}^{N} x(t_i)^2\Big] \tag{1.52}$$

and the *risk costs*

$$\alpha\Sigma \cdot \mathbb{E}\Big[\sum_{i=0}^{N} X(t_i)^2\Big].$$

Intuitively, using a dark pool reduces the overall costs. However, this does not necessarily apply to *both* components of the costs. It is possible that the dark pool reduces one component, while the other one is increased. The following proposition shows that both the total costs and the impact costs of the optimal trading strategy are strictly decreasing in the probability of execution p. On the other hand, the risk costs are not decreasing in $p \in (0,1)$ in general. In particular, it is generally not true that the risk costs of using a dark pool are smaller than the risk costs of not using a dark pool.

Proposition 1.3.9. *Let $(x^*(t_i), y^*(t_i))_{i=0,\dots,N}$ be the optimal strategy and define $X(t_i)$ recursively by*

$$X(t_i) = X(t_{i+1}) - x^*(t_i) - z^*(t_i)$$

as before.

(i) For $i = 1, ..., N-1$, $C(t_i, p)$ is strictly decreasing in p for $p \in (0,1)$.

(ii) If $\alpha\Sigma > 0$, then the risk costs

$$\alpha\Sigma \cdot \mathbb{E}\Big[\sum_{i=0}^{N} X(t_i)^2\Big]$$

are strictly increasing for $p \in (0, \frac{\alpha\Sigma}{\Lambda+\alpha\Sigma})$ and strictly decreasing for $p \in (\frac{\alpha\Sigma}{\Lambda+\alpha\Sigma}, 1)$.

(iii) The impact costs

$$\Lambda \cdot \mathbb{E}\Big[\sum_{i=0}^{N} x^*(t_i)^2\Big]$$

are strictly decreasing for $p \in (0,1)$.

Proof. (i) We prove the first assertion by backward induction using the recursion formula for C given in Equation (1.35):

$$C(t_i, p) = \alpha\Sigma + \frac{(1-p)C(t_{i+1}, p)\Lambda}{\Lambda + (1-p)C(t_{i+1}, p)}. \tag{1.53}$$

This implies $C(t_i, p) > 0$ for $i = 0, \dots, N$ by backward induction; recall that $C(t_N, p) = \Lambda + \alpha\Sigma > 0$ is independent of $p < 1$. Let now $i < N$. By the induction hypothesis, we have

$$\frac{\partial C}{\partial p}(t_{i+1}, p) \le 0.$$

By Equation (1.53), we obtain

$$\frac{\partial C}{\partial p}(t_i, p) = \frac{\partial}{\partial p}\Big(\alpha\Sigma + \frac{(1-p)C(t_{i+1}, p)\Lambda}{\Lambda + (1-p)C(t_{i+1}, p)}\Big)$$

$$= \Lambda^2 \frac{(1-p)\frac{\partial C}{\partial p}(t_{i+1}, p) - C(t_{i+1}, p)}{\big(\Lambda + (1-p)C(t_{i+1}, p)\big)^2}$$
$$< 0$$

as desired.

(ii) Applying the optimal strategy, we obtain

$$\mathbb{E}\big[X(t_i)^2\big] = (1-p)^i X^{\mathrm{ne}}(t_i, p)^2 = \Big(\frac{\sinh(\kappa(p)(N+1-i))}{\sinh(\kappa(p)(N+1))}\Big)^2 X(t_0)^2$$

by Equation (1.36). By Lemma 1.3.8, this term is strictly increasing for $p \in (0, \frac{\alpha\Sigma}{\Lambda+\alpha\Sigma})$ and strictly decreasing for $p \in (\frac{\alpha\Sigma}{\Lambda+\alpha\Sigma}, 1)$.

(iii) Assertions (i) and (ii) imply that the impact costs are strictly decreasing for $p \in (0, \frac{\alpha\Sigma}{\Lambda+\alpha\Sigma})$. We can thus limit our attention to $p \in (\frac{\alpha\Sigma}{\Lambda+\alpha\Sigma}, 1)$, in particular $\kappa'(p) > 0$ by (1.42). We obtain

$$\begin{aligned}\mathbb{E}[x^*(t_i)^2] &= (1-p)^i x^{\mathrm{ne}}(t_i, p)^2 \\ &= \frac{1}{1-p}\Big(\frac{\sqrt{1-p}\,\sinh(\kappa(p)(N+1-i)) - \sinh(\kappa(p)(N-i))}{\sinh(\kappa(p)(N+1))}\Big)^2 X(t_0)^2\end{aligned}$$

by Equation (1.37). For $i = 0, \dots, N$, we set $a = N+1-i$, $b = N+1$ and define the function

$$g_i(p) := \frac{\sinh((a+1)\kappa(p)) - \frac{1}{\sqrt{1-p}}\sinh(a\kappa(p))}{\sinh(b\kappa(p))}.$$

As in the proof of Proposition 1.3.7, we obtain (cf. Inequality (1.48))

$$\sinh(\kappa(p)) \geq \frac{p}{2\sqrt{1-p}}.$$

Using the addition formulae for hyperbolic functions and the definition of $\kappa(p)$, we deduce

$$\begin{aligned}&\sinh((a+1)\kappa(p)) - \frac{1}{\sqrt{1-p}}\sinh(a\kappa(p)) \\ &= \sinh(a\kappa(p))\cosh(\kappa(p)) + \cosh(a\kappa(p))\sinh(\kappa(p)) - \frac{1}{\sqrt{1-p}}\sinh(a\kappa(p)) \\ &= \sinh(a\kappa(p)) \underbrace{\Big(\cosh(\kappa(p)) - \frac{1}{\sqrt{1-p}}\Big)}_{\geq \frac{\sqrt{1-p}}{2}(1+\frac{1}{1-p}) - \frac{1}{\sqrt{1-p}} = \frac{-p}{2\sqrt{1-p}}} + \cosh(a\kappa(p)) \underbrace{\sinh(\kappa(p))}_{\geq \frac{p}{2\sqrt{1-p}}} \\ &\geq \frac{p}{2\sqrt{1-p}}(\cosh(a\kappa(p)) - \sinh(a\kappa(p))) > 0.\end{aligned}$$

To finish the proof, it is thus sufficient to show that

$$\frac{d}{dp}g_i(p) \geq 0 \quad \text{for all } i \quad \text{and} \quad \frac{d}{dp}g_{i_0}(p) > 0 \quad \text{for at least one } i_0.$$

We compute

$$\begin{aligned}\frac{d}{dp}g_i(p) = &-\frac{1}{2(1-p)^{\frac{3}{2}}}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} \\ &+ \frac{\kappa'(p)}{\sqrt{1-p}}\underbrace{\left(\sqrt{1-p}\frac{\partial}{\partial\kappa}\frac{\sinh((a+1)\kappa(p))}{\sinh(b\kappa(p))} - \frac{\partial}{\partial\kappa}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}\right)}_{=:h_i(p)}. \end{aligned} \quad (1.54)$$

Let us first assume that

$$-\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} \leq \frac{\partial}{\partial\kappa}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}.$$

For $i \neq 0$, we have

$$\frac{\partial}{\partial\kappa}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} < 0 \quad \text{and} \quad \frac{\kappa'(p)}{\sqrt{1-p}} < \frac{1}{2(1-p)^{\frac{3}{2}}}$$

by Lemma 1.3.8 (recall $\kappa'(p) > 0$) and therefore

$$\frac{\kappa'(p)}{\sqrt{1-p}}h_i(p) < -\frac{1}{2(1-p)^{\frac{3}{2}}}\frac{\partial}{\partial\kappa}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} \leq \frac{1}{2(1-p)^{\frac{3}{2}}}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))},$$

hence

$$\frac{d}{dp}g_i(p) < 0.$$

For $i = 0$, we have $a = b$,

$$\frac{\partial}{\partial\kappa}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} = 0$$

and thus

$$\frac{d}{dp}g_i(p) \leq 0.$$

We can therefore assume that

$$-\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} > \frac{\partial}{\partial\kappa}\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}$$

from now on. Using the addition formulae for hyperbolic functions, we obtain

$$h_i(p) = \sqrt{1-p}\frac{\partial}{\partial\kappa}\frac{\sinh(a\kappa(p))\cosh(\kappa(p)) + \cosh(a\kappa(p))\sinh(a\kappa(p))}{\sinh(b\kappa(p))}$$

$$
\begin{aligned}
&- \frac{\partial}{\partial \kappa} \frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} \\
&= \sqrt{1-p} \frac{\partial}{\partial \kappa} \Big(\big(\cosh(\kappa(p)) + \sinh(\kappa(p)) - \frac{1}{\sqrt{1-p}} \big) \frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} \\
&\quad + \big(\cosh(a\kappa(p)) - \sinh(a\kappa(p)) \big) \frac{\sinh(\kappa(p))}{\sinh(b\kappa(p))} \Big) \\
&= \sqrt{1-p} \underbrace{\big(\cosh(\kappa(p)) + \sinh(\kappa(p)) - \frac{1}{\sqrt{1-p}} \big)}_{>0} \underbrace{\frac{\partial}{\partial \kappa} \frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}}_{<-\frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}} \\
&\quad + \sqrt{1-p} \big(\cosh(\kappa(p)) + \sinh(\kappa(p)) \big) \frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))} \\
&\quad + \sqrt{1-p} \underbrace{\big(\cosh(a\kappa(p)) - \sinh(a\kappa(p)) \big)}_{>0} \underbrace{\frac{\partial}{\partial \kappa} \frac{\sinh(\kappa(p))}{\sinh(b\kappa(p))}}_{<0 \text{ by Lemma 1.3.8}} \\
&\quad + a\sqrt{1-p} \underbrace{\big(\sinh(a\kappa(p)) - \cosh(a\kappa(p)) \big)}_{<0} \frac{\sinh(\kappa(p))}{\sinh(b\kappa(p))} \\
&< \frac{\sinh(a\kappa(p))}{\sinh(b\kappa(p))}.
\end{aligned}
$$

We note again that

$$0 < \frac{\kappa'(p)}{\sqrt{1-p}} < \frac{1}{2(1-p)^{\frac{3}{2}}}$$

and deduce by Equation (1.54) that

$$\frac{d}{dp} g_i(p) < 0,$$

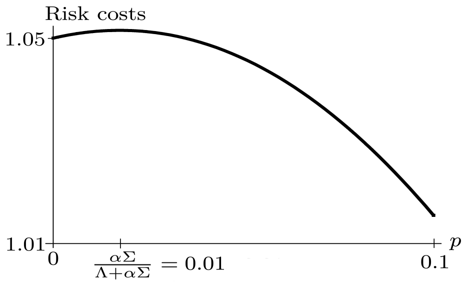
finishing the proof.

□

We illustrate the dependence of the two components of the costs of the optimal strategy on p in Figure 1.5. The left graph shows that the risk costs are increasing for small p while the impact costs are decreasing in p on the whole interval $(0, 1)$ (right graph). Therefore, the reduction of the impact costs outweighs the increase of the risk costs for small p.

1.3.4. Liquidating a portfolio of two assets

If a risk-averse investor has to liquidate a portfolio of multiple assets ($n \geq 2$), then correlation between the assets comes into play. It might no longer be optimal to always place the remaining portfolio into the dark pool. For example, a trader liquidating a well diversified portfolio consisting of two assets will most likely not want to risk losing



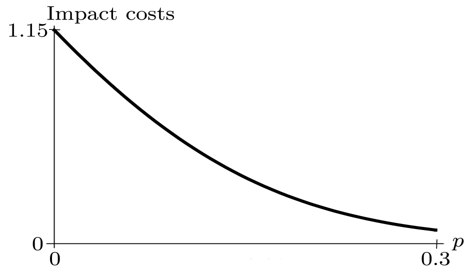


Figure 1.5.: Risk costs (left graph) and impact costs (right graph) of the optimal strategy dependent on the probability of execution in the dark pool p. For $p \in (0, \frac{\alpha\Sigma}{\Lambda+\alpha\Sigma})$, the risk costs are increasing whereas the impact costs are strictly decreasing on the whole interval $(0, 1)$. $N = 15$, $\Lambda = 15$, $\Sigma = \frac{1}{15}$, $\alpha = 4$.

her balanced position by being executed in only one of the two assets. It is therefore important to analyze the dependence of the optimal strategy on the model parameters, especially on the correlation of the assets. Intuitively, we expect the optimal order placement to depend on the correlation of the n assets in the following sense.

If the portfolio is well diversified at the beginning, the orders in the dark pool should be much smaller than the current portfolio as the trader does not want to risk entering an undiversified position. The trading speed in the primary venue should be almost constant since the portfolio position bears little risk and a constant trading speed minimizes the price impact cost (cf. the thin lines in Figure 1.3). If the portfolio is poorly diversified, the orders should initially be comparatively large both in the primary venue and in the dark pool. They might even be larger than the current portfolio for risk mitigating reasons as the execution of the dark pool order for one of the assets can lead to a less risky overall position.

As we are not able to compute the solution of the Optimization Problem (OPT_{dis}) in closed form for $n \geq 2$, we illustrate the above intuitions by the following numerical example.

Example 1.3.10. *We consider different portfolios of two highly correlated stocks with*

$$\Sigma = \frac{1}{500} \cdot \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}.$$

We model the second stock as being more liquid. This is reflected by both a smaller price impact and a higher execution probability in the dark pool[1] *compared to the first asset:*

$$\Lambda = 500 \cdot \begin{pmatrix} 3 & 0 \\ 0 & 0.2 \end{pmatrix},$$

[1]Intuitively, we expect a close connection between liquidity costs in the primary venue and probability of execution in the dark pool. However, we are not aware of any empirical work supporting this.

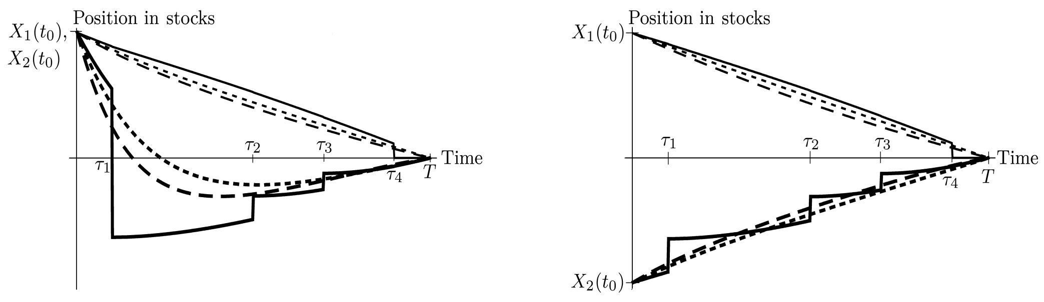


Figure 1.6.: Evolution of a portfolio consisting of two highly correlated stocks over time. The left figure illustrates the poorly diversified portfolio, the right figure the well diversified portfolio. In both pictures thin lines are used for the less liquid first stock and thick lines for the more liquid second stock. Dotted lines correspond to trading without the dark pool, dashed lines correspond to the expected position in the assets if the dark pool is used and solid lines correspond to a realization of the liquidation process using the dark pool, where dark pool orders for the second stock are executed at times τ_1, τ_2, τ_3 and for the first stock only at time τ_4.

$$
\begin{aligned}
\mathbb{P}[\textit{No dark pool execution}] =& \frac{993}{1000}, \\
\mathbb{P}[\textit{Dark pool execution of first asset only}] =& \frac{1}{1000}, \\
\mathbb{P}[\textit{Dark pool execution of second asset only}] =& \frac{6}{1000}, \\
\mathbb{P}[\textit{Dark pool execution of both assets}] =& 0.
\end{aligned}
$$

We consider the following two portfolios in more depth:

(i) Long positions in both stocks, i.e., a poorly diversified portfolio:

$$X(t_0) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

(ii) A long position in the first and a short position in the second stock, i.e., a well diversified portfolio:

$$X(t_0) = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Figure 1.6 shows the evolution of the two portfolios if a risk-averse investor ($\alpha = 4$) applies the optimal strategy. The left picture corresponds to the first case, the right one to the second. In both pictures thin lines are used for the first stock and thick lines for the second. Dotted lines correspond to trading without the dark pool, dashed lines correspond to the expected position in the assets if the dark pool is used and the solid lines correspond to a realization of the liquidation process using the dark pool, where the

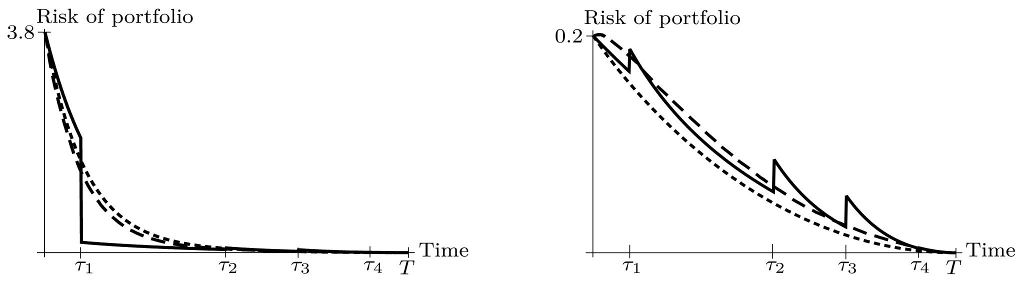


Figure 1.7.: Evolution of risk $X(t_i)^\top \Sigma X(t_i)$ over time for the liquidation paths in Figure 1.6. The dashed line denotes the expected evolution of risk. Note the different scales in the left and the right graph.

dark pool orders for the second stock are executed at times τ_1, τ_2, τ_3 and for the first stock only at time τ_4, i.e., dark pool orders for the more liquid stock are executed several times before any execution in the less liquid stock takes place.

For the poorly diversified portfolio, the trader tries to improve her risky position by trading out of the second stock. For this stock, trading in the primary venue is less expensive and being executed in the dark pool is more probable. If the trader uses the dark pool, this process on average evolves significantly faster than without the dark pool.

For the well diversified portfolio, the portfolio position is decreasing almost linearly in time in all cases. We expect to trade only slightly faster if we use the dark pool. Note that this corresponds to the intuition given at the beginning of the section: It is most profitable to trade out of the position almost evenly.

Additionally, orders in the dark pool are very large for the poorly diversified portfolio and comparatively small for the well diversified portfolio. The reason can be observed in Figure 1.7. The solid lines in the two pictures represent the evolution of portfolio risk $X(t_i)^\top \Sigma X(t_i)$ over time corresponding to the realized liquidation paths in Figure 1.6. The dotted lines represent the evolution of risk if the optimal strategy without the dark pool is used and the dashed lines represent the expected evolution of risk.

As long as the portfolio is poorly diversified, the risk is relatively large and is significantly decreased by a large execution in the dark pool (left picture). However, if it is well diversified as in the right picture, each execution in the dark pool increases the risk. Therefore, the dark pool saves price impact costs but potentially increases risk costs in this case. Note also that in the case of an initially well diversified portfolio, the expected risk can be larger than the risk without using the dark pool.

Our model allows for dependencies between the liquidities of the assets in the dark pool. In Example 1.3.10 we assumed that the probability of simultaneous execution in the dark pools is zero. This is not necessarily the case in reality. Intuitively, we expect the optimal orders in the dark pool to depend strongly on the correlation of the liquidities. For a well diversified portfolio, large probability of simultaneous execution decreases the risk of losing the balanced position by a dark pool execution. Therefore, we expect the absolute value of the optimal orders in the dark pool to increase in the

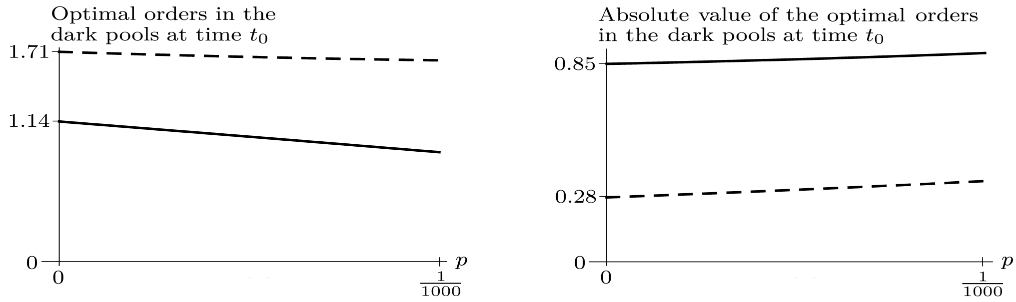


Figure 1.8.: Dependence of the optimal orders in the dark pool on the probability of simultaneous execution. The left picture shows the optimal orders at time t_0 for a poorly diversified portfolio ($X(t_0) = (1,1)^\top$) and the right picture for a well diversified portfolio ($X(t_0) = (1,-1)^\top$). In both pictures the solid line represents the first (illiquid) asset and the dashed line the second (liquid) asset.

probability of simultaneous execution. For a poorly diversified portfolio, large probability of simultaneous execution decreases the likelihood of reaching a more balanced position by a dark pool execution for only one of the two assets, and we therefore expect the optimal orders to decrease in the probability of simultaneous execution. Again, we illustrate these intuitions by a numerical example.

Example 1.3.11. *Let us consider the same stocks as in Example 1.3.10 but this time*

$$\begin{aligned} p := \mathbb{P}[\textit{Dark pool execution of both assets}] &\in [0, \frac{1}{1000}], \\ \mathbb{P}[\textit{Dark pool execution of first asset only}] &= \frac{1}{1000} - p, \\ \mathbb{P}[\textit{Dark pool execution of second asset only}] &= \frac{6}{1000} - p, \\ \mathbb{P}[\textit{No dark pool execution}] &= \frac{993}{1000} + p. \end{aligned}$$

The left picture of Figure 1.8 shows the optimal orders at time t_0 in the dark pool dependent on p for a poorly diversified portfolio ($X(t_0) = (1,1)^\top$) and the right picture shows the absolute value of the optimal orders for a well diversified portfolio ($X(t_0) = (1,-1)^\top$). In both pictures the solid line represents the first (illiquid) asset and the dashed line the second (liquid) asset. Thus, the intuitions preceding the example are supported.

Figure 1.8 also illustrates that the optimal strategy depends strongly on the relative liquidity of the two stocks. Risk mitigation plays a minor role for the illiquid stock as impact costs outweigh risk. Therefore, both in the poorly and in the well diversified case, the orders in the dark pool for the illiquid stock are relatively close to the remainder of the position in the stock, whereas this is not the case for the liquid stock.

1.4. Adverse selection

In Section 1.3 we assumed that the price increments $\epsilon(t_{i+1})$ are independent of the liquidity $a(t_i)$, $b(t_i)$ in the dark pool. This assumption is not always satisfied in reality. For example, other large traders might simultaneously seek liquidity in the dark pool and move prices at the primary trading venue. If price changes $\epsilon(t_{i+1})$ and liquidity $a(t_i), b(t_i)$ in the dark pool at time t_i are correlated, liquidity seeking traders might find that their trades in the dark pool are usually executed just before a favorable price move, i.e., exactly when they *do not* want them to be executed since they miss out on the price improvement. In advance of adverse price movements, they might observe that they rarely find liquidity in the dark pool. We call such a phenomenon "adverse selection".

We keep the assumptions of Section 1.3 except for the fact that we allow for dependencies of dark pool liquidity and future price moves in a *single asset model*. In Section 1.4.1 we specify the market model. As in Section 1.3, we drop Assumption 1.2.1 and can therefore not use the results of Section 1.2. In Section 1.4.2 we prove existence and uniqueness of optimal trading strategies by the same line of reasoning as in the proofs of Proposition 1.2.3 respectively Theorem 1.2.4. As in Section 1.3, $z(t_i)$ depends linearly on $y(t_i)$ by Assumption 1.4.1 (iii). This simplifies the uniqueness proof significantly. Section 1.4.3 is concerned with heuristics about the structure of the value function and the optimal strategy. Using these heuristics, we derive recursions for both in Section 1.4.4. We discuss the properties of the value function and the optimal strategy in Section 1.4.5. These results enable us to derive closed form solutions of the value function and the optimal strategy in Section 1.4.6.

1.4.1. Model description

We only treat the case of *single asset liquidation* ($n = 1$) here. As in Section 1.3, we do not require Assumption 1.2.1 and consider a possibly infinite probability space Ω. We replace Assumption 1.3.1 by the following assumption.

Assumption 1.4.1. *(i) For $i = 0, \dots, N$,*

$$f_i(x(t_i)) := f_i(x(t_0), \dots, x(t_i)) := \Lambda x(t_i)$$

for a positive definite matrix $\Lambda \in \mathbb{R}^{n \times n}$.

(ii) $\tilde{P}$ is an $\mathbb{F}$-martingale and

$$\Sigma(t_i) = \Sigma(t_j) =: \Sigma \quad \text{for all } i, j = 0, \dots, N.$$

(iii) For $i = 0, \dots, N$,

$$a(t_i), b(t_i) \in \{0, \infty\},$$

$((a(t_i), b(t_i)))_{i=0,\dots,N}$ is identically distributed and there exists a nonnegative real number Γ such that for $i = 0, \dots, N$,

$$\mathbb{E}[\epsilon(t_{i+1}) | a(t_i) = \infty] = -\Gamma, \quad \mathbb{E}[\epsilon(t_{i+1}) | b(t_i) = \infty] = \Gamma.$$

(iv) *There exists a real number* $p \in [0,1)$ *such that for all* $i = 0, \dots, N$,

$$p = \mathbb{P}[a(t_i) = \infty] = \mathbb{P}[b_k(t_i) = \infty].$$

The only difference compared to Assumption 1.3.1 is (iii) (recall that for the case $n = 1$, Assumption 1.3.1 (iv) is equivalent to Assumption 1.4.1 (iv)). The model in this section is thus a generalization of the model in Section 1.3 for the case $n = 1$. On the other hand, the model in Section 1.3 generalizes the case $\Gamma = 0$ of this model for general $n \in \mathbb{N}$.

The martingale property (Assumption 1.4.1 (ii)), the dependence of future price moves on dark pool liquidity (Assumption 1.4.1 (iii)), Assumption 1.4.1 (iv) and the liquidation constraint (Definition 1.1.4 (i)) imply

$$\mathbb{E}\Big[\sum_{j=i}^{N}(x(t_j) + z(t_j))(\tilde{P}(t_i) - \tilde{P}(t_j))\Big] = \mathbb{E}\Big[\sum_{j=i}^{N} p\Gamma|y(t_j)|\Big].$$

The Optimization Problem ($\mathrm{OPT}^{\mathrm{gen}}_{\mathrm{dis}}$) thus becomes

$$v(t_i, X(t_i)) = \inf_{(x,y)\in\mathbb{A}(t_i,X(t_i))}\Big\{\mathbb{E}\Big[\sum_{j=i}^{N}\left(\Lambda x(t_j)^2 + p\Gamma|y(t_j)|\right)\Big] + \alpha\mathbb{E}\Big[\sum_{j=i}^{N}\Sigma X(t_j)^2\Big]\Big\}. \quad (\overline{\mathrm{OPT}_{\mathrm{dis}}})$$

Adverse selection is irrelevant for trading without dark pools (or equivalently $p = 0$), so we can assume $p > 0$ from now on.

1.4.2. Existence and uniqueness of optimal strategies

If an optimal strategy at time t_i for an asset position $X \in \mathbb{R}$ exists, we denote it by

$$(x^*(t_i, X), y^*(t_i, X)) \in \mathbb{R} \times \mathbb{R}.$$

Theorem 1.4.2. *Let* $i = 0, \dots, N$.

(i) $v(t_i, \cdot)$ *is symmetric and monotone in the following sense:*

$$v(t_i, -X(t_i)) = v(t_i, X(t_i)) \quad \text{and}$$
$$|X(t_i)| \le |Y(t_i)| \Rightarrow v(t_i, X(t_i)) \le v(t_i, Y(t_i)).$$

(ii) $v(t_i, \cdot)$ *is strictly convex, and for* $X(t_i) \in \mathbb{R}$ *there exits a unique optimal trading strategy* $(x^*, y^*) \in \mathbb{A}(t_i, X(t_i))$ *realizing the minimum in Equation* ($\overline{\mathrm{OPT}_{\mathrm{dis}}}$).

Proof. The Optimization Problem ($\overline{\mathrm{OPT}_{\mathrm{dis}}}$) yields the Bellman equation

$$\begin{aligned} v(t_i, X(t_i)) = \inf_{(x,y)\in\mathbb{R}\times\mathbb{R}} \{&\Lambda x^2 + \alpha\Sigma X(t_i)^2 + p|y|\Gamma + pv(t_{i+1}, X(t_i) - x - y) \\ &+ (1-p)v(t_{i+1}, X(t_i) - x)\}. \end{aligned} \quad (1.55)$$

We define the function $\tilde{v}(X(t_i), \cdot) : \mathbb{R}^2 \longrightarrow \mathbb{R}$ by

$$\begin{aligned}\tilde{v}(X(t_i), x, y) := \Lambda x^2 + \alpha \Sigma X(t_i)^2 + p|y|\Gamma + pv(t_{i+1}, X(t_i) - x - y) \\ + (1-p)v(t_{i+1}, X(t_i) - x).\end{aligned} \tag{1.56}$$

(i) Symmetry can be shown by a straightforward backward induction on $i \leq N$. We also prove monotonicity by backward induction. The case $i = N$ is clear, and we set $i < N$. We can assume without loss of generality that $0 \leq X(t_i) \leq Y(t_i)$ by symmetry. For any admissible pair of orders $(x(t_i), y(t_i)) \in \mathbb{R}^2$ we define

$$\tilde{x}(t_i) := \begin{cases} x(t_i) & \text{if } X(t_i) - x(t_i) \geq 0 \\ X(t_i) & \text{else,} \end{cases}$$

$$\tilde{y}(t_i) := \begin{cases} y(t_i) & \text{if } X(t_i) - x(t_i), X(t_i) - x(t_i) - y(t_i) \geq 0 \\ X(t_i) - x(t_i) & \text{if } X(t_i) - x(t_i) \geq 0,\ X(t_i) - x(t_i) - y(t_i) < 0 \\ 0 & \text{else.} \end{cases}$$

In all three possible cases a simple argument using the induction hypothesis establishes

$$\tilde{v}(X(t_i), \tilde{x}(t_i), \tilde{y}(t_i)) \leq \tilde{v}(Y(t_i), x(t_i), y(t_i)),$$

finishing the proof.

(ii) The strict convexity of $v(t_i, \cdot)$ and the existence and uniqueness of an optimal trading strategy can also be established by backward induction. Again, the case $i = N$ is clear. Let therefore $i < N$ and $X(t_i) \in \mathbb{R}$. Note that by the induction hypothesis and (i), $\tilde{v}(X(t_i), \cdot)$ is strictly convex (in particular continuous) and

$$\lim_{\|(x,y)\| \to \infty} \tilde{v}(X(t_i), x, y) = \infty.$$

Therefore, a unique optimal strategy exists. Strict convexity of $v(X(t_i), \cdot)$ follows by the strict convexity of $\tilde{v}(t_i, \cdot)$ using the Bellman Equation (1.55).

□

We can deduce properties of the optimal trading strategy from Theorem 1.4.2 that are helpful for obtaining recursions for the value function and the optimal strategy in Section 1.4.4 (cf. the proof of Theorem 1.4.4). The first assertion of the next corollary means that neither changing the direction of trading nor changing the sign of the position can be optimal.

Corollary 1.4.3. *Let $i = 0, \ldots, N$.*

(i) For $X(t_i) > 0$, we have

$$x^*(t_i, X(t_i)), y^*(t_i, X(t_i)) \geq 0 \quad \textit{and} \quad x^*(t_i, X(t_i)) + y^*(t_i, X(t_i)) \leq X(t_i). \tag{1.57}$$

(ii) $(x^*(t_i,\cdot), y^*(t_i,\cdot))$ *is continuous.*

Proof. (i) Let $(x(t_i), y(t_i)) \in \mathbb{R}^2$ be any admissible pair of orders not satisfying (1.57). We define

$$\tilde{x}(t_i) := \begin{cases} x(t_i) & \text{if } x(t_i) \geq 0,\ x(t_i, X(t_i)) \leq X(t_i) \\ X(t_i) & \text{if } x(t_i) \geq 0,\ x(t_i, X(t_i)) > X(t_i) \\ 0 & \text{else,} \end{cases}$$

$$\tilde{y}(t_i) := \begin{cases} y(t_i) & \text{if } 0 \leq y(t_i) \leq X(t_i) - \tilde{x}(t_i) \\ X(t_i) - \tilde{x}(t_i) & \text{if } y(t_i) > X(t_i) - \tilde{x}(t_i) \\ 0 & \text{else.} \end{cases}$$

For $\tilde{v}$ as in Equation (1.56) the Bellman Equation (1.55) and Theorem 1.4.2 (i) yield

$$\tilde{v}(X(t_i), \tilde{x}(t_i), \tilde{y}(t_i)) < \tilde{v}(Y(t_i), x(t_i), y(t_i))$$

for all possible cases, finishing the proof.

(ii) Let $X \in \mathbb{R}$ and consider a sequence of real numbers $(X_k)_k$ such that

$$\lim_{k\to\infty} X_k = X \in \mathbb{R}.$$

By (i), the sequences $(x^*(t_i, X_k))_k$ and $(y^*(t_i, X_k))_k$ are bounded and therefore there exist convergent subsequences $(x^*(t_i, X_{k_l}))_l$, $(y^*(t_i, X_{k_l}))_l$ such that

$$\lim_{l\to\infty} x^*(t_i, X_{k_l}) = x_0 \in \mathbb{R}, \quad \lim_{l\to\infty} y^*(t_i, X_{k_l}) = y_0 \in \mathbb{R}.$$

Assume that $(x_0, y_0) \neq (x^*(t_i, X), y^*(t_i, X))$. By the uniqueness of the optimal trading strategy and the continuity of $v(t_i, \cdot)$ and $\tilde{v}(\cdot)$ as in Equation (1.56), we obtain

$$\begin{aligned} \tilde{v}(X, x_0, y_0) &> \tilde{v}(X, x^*(t_i, X), y^*(t_i, X)) = v(t_i, X) = \lim_{l\to\infty} v(t_i, X_{k_l}) \\ &= \lim_{l\to\infty} \tilde{v}(X_{k_l}, x^*(t_i, X_{k_l}), y^*(t_i, X_{k_l})) = \tilde{v}(X, x_0, y_0), \end{aligned}$$

a contradiction.

□

1.4.3. Heuristics for the value function and the optimal strategy

So far we proved existence and uniqueness of the optimal trading strategy. The aim of this section is to derive the structure of the value function and the optimal strategy heuristically. In Section 1.4.4 we formalize these heuristics and present a rigorous proof (Theorem 1.4.4). The proof of Theorem 1.4.4 is constructive and based on the heuristic considerations of this section.

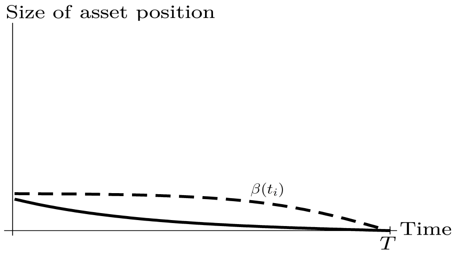


Figure 1.9.: Small asset positions $X(t_i) \leq \beta(t_i)$ are expected to stay below $\beta(t_j)$ for all $j \geq i$ if the candidate optimal strategy in Equation (1.9) is applied.

Let $i < N$ and consider a nonnegative asset position $X(t_i) \geq 0$. By Corollary 1.4.3 (i), the Bellman Equation (1.55) becomes

$$v(t_i, X(t_i)) = \min_{x,y \geq 0} \{\Lambda x^2 + \alpha \Sigma X(t_i)^2 + p\Gamma y + pv(t_{i+1}, X(t_i) - x - y) \\ + (1-p)v(t_{i+1}, X(t_i) - x)\}. \tag{1.58}$$

For $\Gamma = 0$, we obtain the single asset case of Section 1.3.3. In this case the value function is quadratic, and the optimal strategy is linear in the asset position. Because of the linear term

$$p\Gamma y$$

we cannot expect this to be the case for general $\Gamma > 0$. We rather expect the value function to be a quadratic polynomial:

$$v(t_i, X(t_i)) = \bar{C}_1(t_i)X(t_i)^2 + \bar{C}_2(t_i)X(t_i) + \bar{C}_3(t_i). \tag{1.59}$$

It turns out that this ansatz does not capture the whole complexity of the Optimization Problem ($\overline{\text{OPT}_{\text{dis}}}$). However, it is a useful starting point for the following considerations.

The term $p\Gamma y$ in the Bellman Equation (1.58) is equivalent to assuming that orders in the dark pool yield linear costs. For large orders the quadratic costs of trading in the primary venue outweigh the linear costs of trading in the dark pool. However, if a trader aims to liquidate a sufficiently small asset position, then dark pool orders can be more expensive then trading at the primary venue. For this reason, we expect that there exists a time-dependent boundary $\beta(t_i)$ ($i = 0, \ldots, N$) such that for

$$X(t_i) \leq \beta(t_i),$$

the optimal dark pool order is

$$y^*(t_i, X(t_i)) = 0$$

and the value function and the optimal strategy in the primary venue are the same as without dark pool as given in Section 1.3.3:

$$v(t_i, X(t_i)) = C(t_i, 0)X(t_i)^2,$$

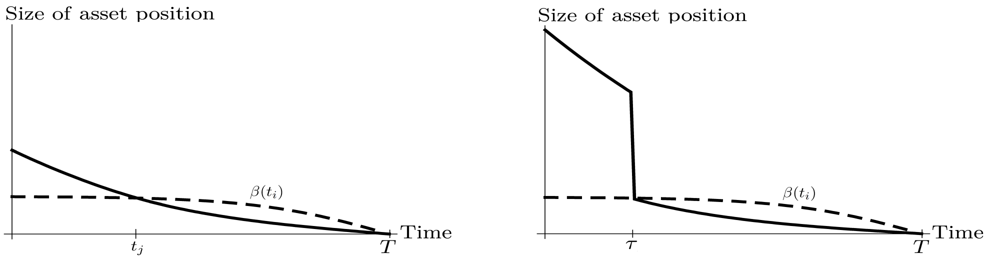


Figure 1.10.: Asset positions $X(t_i) > \beta(t_i)$. The position in the left picture is sufficiently small for the optimal asset position to cross the boundary at time t_j. The larger position in the right picture is sufficiently large such that the optimal asset position only crosses the boundary if the optimal dark pool order is executed (which happens at time τ in the displayed scenario).

$$x^*(t_i, X(t_i)) = A(t_i, 0)X(t_i). \tag{1.60}$$

If the asset position is below the boundary at time t_i, we expect it to stay below it for the remainder of the trading horizon $[t_i, T]$ if the candidate optimal trading strategy given by Equation (1.60) is applied (see Figure 1.9 for an illustration).

Following the above considerations, we expect at least two different trading regions

$$[0, \beta(t_i)] \quad \text{and} \quad (\beta(t_i), \infty).$$

In the first one the optimal dark pool order should be zero and in the second one it should be greater than zero. Given an asset position

$$X(t_i) > \beta(t_i),$$

there should again be roughly two possibilities. Either the order is sufficiently small such that applying the optimal strategy in the primary venue, the optimal asset position crosses the boundary $\beta(t_j)$ at some later point in time t_j, $j = i+1, \dots, N-1$ without any execution in the dark pool. Or it is sufficiently large such that the optimal asset position without any dark pool execution never crosses the boundary. We illustrate the first case in the left picture and the second case in the right picture of Figure 1.10. In both cases we expect that optimal orders in the dark pool are sufficiently large such that the optimal position crosses the boundary immediately after the execution of a dark pool order.

There are $N+1-i$ possible points in time where the optimal asset position can cross the boundary β. This suggests that we should distinguish between $N+1-i$ rather than two trading regions, dependent on the time t_j ($j = i+1, \dots, N$) the initial asset position $X(t_i)$ crosses the boundary $\beta(t_j)$ without an execution in the dark pool. More precisely, we expect that there exists a function

$$\bar{X}(t_i, \cdot) : \{t_i, \dots, t_{N+1}\} \to [0, \infty],$$

which divides the positive real line into trading regimes

$$0 =: \bar{X}(t_i, t_i) < \beta(t_i) =: \bar{X}(t_i, t_{i+1}) < \cdots < \bar{X}(t_i, t_{N+1}) := \infty$$

such that

(i) for
$$X(t_i) \in [\bar{X}(t_i, t_i), \bar{X}(t_i, t_{i+1})] = [0, \beta(t_i)],$$
the dark pool is not used,

(ii) for $j = i+1, \dots, N-1$,
$$X(t_i) \in (\bar{X}(t_i, t_j), \bar{X}(t_i, t_{j+1})],$$
the dark pool is only used in the first $j-i$ trading periods, and the optimal trading strategy crosses the boundary $\beta(t_j) =: \bar{X}(t_j, t_{j+1})$ at time t_j if no dark pool order has been executed before,

(iii) for
$$X(t_i) \in (\bar{X}(t_i, t_N), \bar{X}(t_i, t_{N+1})) = (\bar{X}(t_i, t_N), \infty),$$
the dark pool is used at all trading times $t_i, \dots, t_{N-1}$ until execution.

Let us now return to the beginning of the section (cf. Equation (1.59)). We expect that in each of the $N+1-i$ trading regions the value function is a quadratic polynomial with the coefficients depending on the trading regime:

$$v(t_i, X(t_i)) = C_1(t_i, t_j)X(t_i)^2 + C_2(t_i, t_j)X(t_i) + C_3(t_i, t_j)$$

for $X(t_i) \in (\bar{X}(t_i, t_j), \bar{X}(t_i, t_{j+1}))$. This proves to be the correct ansatz, and we are able to obtain backward recursions for the coefficients C_1, C_2, C_3 and the functions $\bar{X}$ in Theorem 1.4.4.

1.4.4. Structure of the optimal strategy and the value function

We confirm the intuitions about trading regions and the form of the value function given in Section 1.4.3 in Theorem 1.4.4. By backward induction, we obtain recursions both for the value function and for the optimal strategy. We consider the case

$$\alpha\Sigma > 0$$

of a risk-averse investor first. The simpler case $\alpha\Sigma = 0$ is treated at the end of this section.

Theorem 1.4.4. *The value function $v(t_i, \cdot)$ of the Optimization Problem ($\overline{\text{OPT}_{\text{dis}}}$) is a piecewise quadratic polynomial and the optimal strategy $(x^*(t_i, \cdot), y^*(t_i, \cdot))$ at time t_i is piecewise affine linear.*

More precisely, let $X(t_i)$ be the asset position at time t_i.

(i) There exist

$$0 =: \bar{X}(t_i, t_i) < \cdots < \bar{X}(t_i, t_{N+1}) := \infty$$

such that for

$$X(t_i) \in [\bar{X}(t_i, t_j), \bar{X}(t_i, t_{j+1})),$$

the unique optimal strategy and the value function are given by

$$x^*(t_i, X(t_i)) = A_1(t_i, t_j)X(t_i) + A_2(t_i, t_j), \tag{1.61}$$

$$y^*(t_i, X(t_i)) = B_1(t_i, t_j)X(t_i) + B_2(t_i, t_j), \tag{1.62}$$

$$v(t_i, X(t_i)) = C_1(t_i, t_j)X(t_i)^2 + C_2(t_i, t_j)X(t_i) + C_3(t_i, t_j). \tag{1.63}$$

The coefficients $A_1, \dots, C_3$ and $\bar{X}$ are given recursively in Appendix A.1.

(ii) For $X(t_i) < 0$, the optimal strategy and the value function are given by

$$\begin{aligned} x^*(t_i, X(t_i)) &= -x^*(t_i, -X(t_i)), \\ y^*(t_i, X(t_i)) &= -y^*(t_i, -X(t_i)), \\ v(t_i, X(t_i)) &= v(t_i, -X(t_i)). \end{aligned}$$

Proof. We prove (i). Assertion (ii) follows directly by Theorem 1.4.2 (i). Let therefore $i = 0, \dots N$ and $X(t_i) \geq 0$. By Corollary 1.4.3 (i), we know that the optimal asset position at time t_j fulfills $X(t_j) \geq 0$ for all $j \geq i$.

We proceed by backward induction. For $i = N$ there is only one admissible strategy and all the properties are satisfied.

For the induction step we assume that all properties are valid for time t_{i+1} ($i < N$). By Corollary 1.4.3 (i), the Bellman Equation (1.55) becomes

$$\begin{aligned} v(t_i, X(t_i)) = \min_{x,y \geq 0} \Big\{ &\Lambda x^2 + \alpha \Sigma X(t_i)^2 + py\Gamma \\ &+ pv(t_{i+1}, X(t_i) - x - y) + (1-p)v(t_{i+1}, X(t_i) - x) \Big\}. \end{aligned}$$

(i) As a first step we prove the assertion for $j = i$, i.e., the case where the initial asset position is below the boundary and we expect that the use of the dark pool is not optimal:

$$0 \leq X(t_i) < \bar{X}(t_i, t_{i+1}).$$

The recursion formula for $C(t_i, 0)$, Equation (1.35), implies

$$C(t_{i+1}, 0) < C(t_{i+2}, 0)$$

and therefore we know by Equation (A.10) that

$$\bar{X}(t_{i+1}, t_{i+2}) = \frac{\Gamma(C(t_{i+2}, 0) + \Lambda)}{2C(t_{i+2}, 0)\Lambda} < \frac{\Gamma(C(t_{i+1}, 0) + \Lambda)}{2C(t_{i+1}, 0)\Lambda} = \bar{X}(t_i, t_{i+1}).$$

Assume for now that

$$0 \leq X(t_i) < \bar{X}(t_{i+1}, t_{i+2}) < \bar{X}(t_i, t_{i+1}).$$

This implies

$$0 \leq X(t_i) - x^*(t_i, X(t_i)) - y^*(t_i, X(t_i)) \leq X(t_i) - x^*(t_i, X(t_i)) < \bar{X}(t_{i+1}, t_{i+2})$$

by Corollary 1.4.3 (i). We define the set

$$\bar{A}(X(t_i)) := \{(x, y) \in \mathbb{R}^2 | x, y \geq 0,\ X(t_i) - x - y \geq 0\}$$

and obtain by the backward induction hypothesis that

$$\begin{aligned}
v(t_i, X(t_i)) = \min_{(x,y)\in\bar{A}(X(t_i))} \Big\{ & \Lambda x^2 + \alpha\Sigma X(t_i)^2 + py\Gamma \\
& + p\big(C_1(t_{i+1}, t_{i+1})(X(t_i) - x - y)^2 \\
& \quad + C_2(t_{i+1}, t_{i+1})(X(t_i) - x - y) + C_3(t_{i+1}, t_{i+1})\big) \\
& + (1-p)\big(C_1(t_{i+1}, t_{i+1})(X(t_i) - x)^2 \\
& \quad + C_2(t_{i+1}, t_{i+1})(X(t_i) - x) + C_3(t_{i+1}, t_{i+1})\big)\Big\} \\
\overset{(A.1)}{=} \min_{(x,y)\in\bar{A}(X(t_i))} \Big\{ & \Lambda x^2 + \alpha\Sigma X(t_i)^2 + py\Gamma + pC(t_{i+1}, 0)(X(t_i) - x - y)^2 \\
& + (1-p)C(t_{i+1}, 0)(X(t_i) - x)^2\Big\}. \qquad (1.64)
\end{aligned}$$

Let $\bar{v}(X(t_i), \cdot) : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be given by

$$\begin{aligned}
\bar{v}(X(t_i), x, y) := {} & \Lambda x^2 + \alpha\Sigma X(t_i)^2 + py\Gamma + pC(t_{i+1}, 0)(X(t_i) - x - y)^2 \\
& + (1-p)C(t_{i+1}, 0)(X(t_i) - x)^2.
\end{aligned}$$

In order to minimize $\bar{v}(X(t_i), \cdot)$, we consider the first order condition for minimality

$$\frac{\partial}{\partial x}\bar{v}(X(t_i), x, y) = 0, \quad \frac{\partial}{\partial y}\bar{v}(X(t_i), x, y) = 0.$$

This system of linear equations has the unique solution $(\bar{x}_i, \bar{y}_i)$ given by

$$\bar{x}_i = \frac{C(t_{i+1}, 0)(1-p)}{C(t_{i+1}, 0)(1-p) + \Lambda} X(t_i) + \frac{\Gamma p}{2(C(t_{i+1}, 0)(1-p) + \Lambda)}, \qquad (1.65)$$

$$\bar{y}_i = \frac{\Lambda}{C(t_{i+1}, 0)(1-p) + \Lambda} X(t_i) - \frac{\Gamma(C(t_{i+1}, 0) + \Lambda)}{2C(t_{i+1}, 0)(C(t_{i+1}, 0)(1-p) + \Lambda)}. \qquad (1.66)$$

An elementary calculation shows that the Hessian of $\bar{v}(X(t_i), \cdot)$ is positive definite.

Thus, $(\bar{x}_i, \bar{y}_i)$ is the global minimum of $\bar{v}(X(t_i), \cdot)$. However,

$$\begin{aligned}\bar{y}_i &< \frac{\Lambda}{C(t_{i+1},0)(1-p)+\Lambda}\bar{X}(t_{i+1},t_{i+2}) - \frac{\Gamma(C(t_{i+1},0)+\Lambda)}{2C(t_{i+1},0)(C(t_{i+1},0)(1-p)+\Lambda)} \\ &\overset{(A.10)}{=} 0.\end{aligned}$$

Therefore, $\bar{v}(X(t_i), \cdot)$ is minimal for $y = 0$ on $\bar{A}(X(t_i))$. Hence $y^*(t_i, X(t_i)) = 0$, and $x^*(t_i, X(t_i))$ can be obtained by minimizing

$$\Lambda x^2 + \alpha\Sigma X(t_i)^2 + C(t_{i+1},0)(X(t_i)-x)^2. \tag{1.67}$$

Thus,

$$x^*(t_i, X(t_i)) = \frac{C(t_{i+1},0)}{C(t_{i+1},0)+\Lambda}X(t_i) = A(t_i,0)X(t_i).$$

As $X(t_i) - x^*(t_i, X(t_i))$ and $y^*(t_i, X(t_i))$ are continuous in $X(t_i)$ (Corollary 1.4.3 (ii)),

$$(A(t_i,0)X(t_i), 0)$$

is the optimal strategy at time t_i as long as

$$\bar{y}_i < 0 \quad \text{and} \quad X(t_i) - A(t_i,0)X(t_i) < \bar{X}(t_{i+1}, t_{i+2}),$$

which is equivalent to

$$0 \le X(t_i) < \bar{X}(t_i, t_{i+1}).$$

Plugging this into (1.67), we obtain

$$v(t_i, X(t_i)) = \frac{\alpha\Sigma\Lambda + C(t_{i+1},0)(\Lambda+\alpha\Sigma)}{C(t_{i+1},0)+\Lambda}X(t_i)^2 = C(t_i,0)X(t_i)^2,$$

completing the proof for $j = i$.

We will now show the assertion for $j = i+1, \ldots, N$ by forward induction on j.

(ii) For the induction basis $j = i+1$ let

$$X(t_i) \in [\bar{X}(t_i, t_{i+1}), \bar{X}(t_i, t_{i+2})).$$

For $X(t_i) = \bar{X}(t_i, t_{i+1})$, the case $j = i$ proven above and the continuity of the optimal strategy imply

$$\begin{aligned}X(t_i) - x^*(t_i, X(t_i)) &= X(t_i) - A(t_i,0)X(t_i) \\ &\overset{(1.34),(A.10)}{=} \frac{\Gamma}{2C(t_{i+1},0)} \\ &< \frac{\Gamma(C(t_{i+2},0)+\Lambda)}{2C(t_{i+2},0)\Lambda}\end{aligned} \tag{1.68}$$

$$\stackrel{(A.10)}{=} \bar{X}(t_{i+1}, t_{i+2}),$$

where Inequality (1.68) follows from (1.35) as $\alpha\Sigma > 0$:

$$\frac{\Gamma}{2C(t_{i+1},0)} = \frac{\Gamma(\Lambda + C(t_{i+2},0))}{2(\alpha\Sigma\Lambda + C(t_{i+2},0)(\Lambda + \alpha\Sigma))} < \frac{\Gamma(C(t_{i+2},0) + \Lambda)}{2C(t_{i+2},0)\Lambda}.$$

Again by continuity of the optimal strategy, we obtain that the Bellman equation is given by (1.64) in some neighborhood $(\bar{X}(t_i, t_{i+1}), \bar{X}(t_i, t_{i+1}) + \epsilon)$. We consider the first order conditions of optimality again and obtain that the global minimum $(\bar{x}_i, \bar{y}_i)$ of $\bar{v}(X(t_i), \cdot)$ is as given in Equations (1.65) and (1.66) by the positive definiteness of the Hessian as before. This time we have $\bar{x}_i > 0$, $\bar{y}_i \geq 0$, and by Equations (A.6) - (A.9), $\bar{x}_i$ and $\bar{y}_i$ are of the required form:

$$\begin{aligned} \bar{x}_i &= A_1(t_i, t_{i+1})X(t_i) + A_2(t_i, t_{i+1}), \\ \bar{y}_i &= B_1(t_i, t_{i+1})X(t_i) + B_2(t_i, t_{i+1}). \end{aligned}$$

Using the continuity of $(x^*(t_i, \cdot), y^*(t_i, \cdot))$ again, we deduce that $(\bar{x}_i, \bar{y}_i)$ is the optimal strategy as long as

$$\bar{y}_i \geq 0 \quad \text{and} \quad X(t_i) - \bar{x}_i < \bar{X}(t_{i+1}, t_{i+2}). \tag{1.69}$$

A straightforward calculation shows that (1.69) is equivalent to

$$X(t_i) \geq \frac{\Gamma(C(t_{i+1},0) + \Lambda)}{2C(t_{i+1},0)\Lambda} = \bar{X}(t_i, t_{i+1}) \quad \text{and} \quad X(t_i) < \bar{X}(t_i, t_{i+2}).$$

Plugging $(\bar{x}_i, \bar{y}_i)$ into $\bar{v}(X(t_i), \cdot)$, we obtain

$$v(t_i, X(t_i)) = C_1(t_i, t_{i+1})X(t_i)^2 + C_2(t_i, t_{i+1})X(t_i) + C_3(t_i, t_{i+1}),$$

with C_1, C_2 and C_3 according to Appendix A.1, which finishes the proof of the induction basis for the forward induction.

(iii) For the induction step, we assume that the assertion is true for some $j = i+1, \ldots N-1$ and let

$$X(t_i) \in [\bar{X}(t_i, t_{j+1}), \bar{X}(t_i, t_{j+2})).$$

By the induction hypothesis, the continuity of $x^*(t_i, \cdot)$ and $y^*(t_i, \cdot)$ and the recursions in Appendix A.1, we have

$$\begin{aligned} &\bar{X}(t_i, t_{j+1}) - x^*(t_i, \bar{X}(t_i, t_{j+1})) \\ &\qquad = \bar{X}(t_i, t_{j+1}) - A_1(t_i, t_j)\bar{X}(t_i, t_{j+1}) - A_2(t_i, t_j) \\ &\qquad \stackrel{(A.6),(A.8)}{=} B_1(t_i, t_j)\bar{X}(t_i, t_{j+1}) - A_2(t_i, t_j) \end{aligned}$$

$$
\begin{aligned}
&\overset{(A.11)}{=} \bar{X}(t_{i+1}, t_{j+1}), \\
\bar{X}(t_i, t_{j+1}) - x^*(t_i, \bar{X}(t_i, t_{j+1})) - y^*(t_i, \bar{X}(t_i, t_{j+1}))& \\
&= \bar{X}(t_i, t_{j+1}) - A_1(t_i, t_j)\bar{X}(t_i, t_{j+1}) - A_2(t_i, t_j) \\
&\quad - B_1(t_i, t_j)\bar{X}(t_i, t_{j+1}) - B_2(t_i, t_j) \\
&\overset{(A.6),(A.8)}{=} -A_2(t_i, t_j) - B_2(t_i, t_j) \\
&\overset{(A.7),(A.9)}{=} \frac{\Gamma}{2C(t_{i+1}, 0)}
\end{aligned}
$$

and

$$
y^*(t_i, \bar{X}(t_i, t_{j+1})) > 0.
$$

Thus, again by continuity of $x^*(t_i, \cdot)$ and $y^*(t_i, \cdot)$, there exists an $\epsilon > 0$ such that for $X(t_i) \in (\bar{X}(t_i, t_{j+1}), \bar{X}(t_i, t_{j+1}) + \epsilon)$,

$$
\begin{aligned}
y^*(t_i, X(t_i)) &> 0, \\
X(t_i) - x^*(t_i, X(t_i)) &< \bar{X}(t_{i+1}, t_{j+2}), && (1.70) \\
X(t_i) - x^*(t_i, X(t_i)) &> \bar{X}(t_{i+1}, t_j), && (1.71) \\
X(t_i) - x^*(t_i, X(t_i)) - y^*(t_i, X(t_i)) &< \bar{X}(t_{i+1}, t_{i+2}).
\end{aligned}
$$

In order to apply the backward induction hypothesis, we want to ensure that the optimal asset position at time t_{i+1}, provided the dark pool order has not been executed, fulfills

$$
X(t_i) - x^*(t_i, X(t_i)) \in [\bar{X}(t_{i+1}, t_{j+1}), \bar{X}(t_{i+1}, t_{j+2}));
$$

notice that Inequalities (1.70) and (1.71) only imply

$$
X(t_i) - x^*(t_i, X(t_i)) \in (\bar{X}(t_{i+1}, t_j), \bar{X}(t_{i+1}, t_{j+2})).
$$

To this end, assume that $X(t_i) - x^*(t_i, X(t_i)) < \bar{X}(t_{i+1}, t_{j+1})$. For this case the backward induction hypothesis would imply

$$
\begin{aligned}
&v(t_i, X(t_i)) \\
&\quad = \min_{\check{A}(X(t_i))} \Big\{ \Lambda x^2 + \alpha\Sigma X(t_i)^2 + py\Gamma + pC(t_{i+1}, 0)(X(t_i) - x - y)^2 \\
&\qquad + (1-p)\big(C_1(t_{i+1}, t_j)(X(t_i)-x)^2 + C_2(t_{i+1}, t_j)(X(t_i)-x) + C_3(t_{i+1}, t_j)\big)\Big\},
\end{aligned}
$$

where

$$
\begin{aligned}
\check{A}(X(t_i)) := \{(x, y) \in \mathbb{R}^2 | x, y \geq 0,\ &\bar{X}(t_{i+1}, t_j) \leq X(t_i) - x \leq \bar{X}(t_{i+1}, t_{j+1}), \\
&0 \leq X(t_i) - x - y \leq \bar{X}(t_{i+1}, t_{i+2})\}.
\end{aligned}
$$

A simple calculation shows that this minimization problem has the unique solution

$$\check{x}_i = A_1(t_i, t_j)X(t_i) + A_2(t_i, t_j),$$
$$\check{y}_i = B_1(t_i, t_j)X(t_i) + B_2(t_i, t_j)$$

for $A_1, \ldots, B_2$ according to Appendix A.1. This implies

$$\begin{aligned} X(t_i) - \check{x}_i &\overset{(A.6),(A.8)}{=} B_1(t_i, t_j)X(t_i) - A_2(t_i, t_j) \\ &> B_1(t_i, t_j)\bar{X}(t_{i+1}, t_j) - A_2(t_i, t_j) \\ &\overset{(A.11)}{=} \bar{X}(t_{i+1}, t_{j+1}), \end{aligned}$$

a contradiction. Therefore,

$$X(t_i) - x(t_i, X(t_i)) \geq \bar{X}(t_{i+1}, t_{j+1}) \quad \text{for} \quad X(t_i) \in (\bar{X}(t_i, t_{j+1}), \bar{X}(t_i, t_{j+1}) + \epsilon).$$

By the backward induction hypothesis, the Bellman Equation (1.55) becomes

$$\begin{aligned} &v(t_i, X(t_i)) \\ &\quad = \min_{\hat{A}(X(t_i))} \Big\{ \Lambda x^2 + \alpha \Sigma X(t_i)^2 + py\Gamma + pC(t_{i+1}, 0)(X(t_i) - x - y)^2 + (1-p) \\ &\qquad (C_1(t_{i+1}, t_{j+1})(X(t_i) - x)^2 + C_2(t_{i+1}, t_{j+1})(X(t_i) - x) + C_3(t_{i+1}, t_{j+1})) \Big\}. \\ &\quad =: \min_{\hat{A}(X(t_i))} \hat{v}(X(t_i), x, y), \end{aligned}$$

where

$$\begin{aligned} \hat{A}(X(t_i)) := \{ &(x, y) \in \mathbb{R}^2 | x, y \geq 0,\ \bar{X}(t_{i+1}, t_{j+1}) \leq X(t_i) - x \leq \bar{X}(t_{i+1}, t_{j+2}), \\ &0 \leq X(t_i) - x - y \leq \bar{X}(t_{i+1}, t_{i+2}) \}. \end{aligned}$$

Again, the Hessian of $\hat{v}(X(t_i), \cdot)$ is positive definite and the gradient equals zero for $(\hat{x}_i, \hat{y}_i)$ such that

$$\begin{aligned} \hat{x}_i &= \frac{C_1(t_{i+1}, t_{j+1})(1-p)}{C_1(t_{i+1}, t_{j+1})(1-p) + \Lambda} X(t_i) + \frac{C_2(t_{i+1}, t_{j+1})(1-p) + \Gamma p}{2(C_1(t_{i+1}, t_{j+1})(1-p) + \Lambda)} \\ &\overset{(A.6),(A.7)}{=} A_1(t_i, t_{j+1})X(t_i) + A_2(t_i, t_{j+1}), \\ \hat{y}_i &= \frac{\Lambda}{C_1(t_{i+1}, t_{j+1})(1-p) + \Lambda} X(t_i) - \\ &\quad \frac{C(t_{i+1}, 0)C_2(t_{i+1}, t_{j+1})(1-p) + \Gamma(C_1(t_{i+1}, t_{j+1})(1-p) + C(t_{i+1}, 0)p + \Lambda)}{2C(t_{i+1}, 0)(C_1(t_{i+1}, t_{j+1})(1-p) + \Lambda)} \\ &\overset{(A.8),(A.9)}{=} B_1(t_i, t_{j+1})X(t_i) + B_2(t_i, t_{j+1}), \end{aligned}$$

i.e., $\hat{x}_i$ and $\hat{y}_i$ are of the required form. Again, by continuity of $(x^*(t_i, \cdot), y^*(t_i, \cdot))$,

$(\hat{x}_i, \hat{y}_i)$ defines the optimal strategy as long as

$$\bar{X}(t_{i+1}, t_{j+1}) \leq X(t_i) - \hat{x}_i \leq \bar{X}(t_{i+1}, t_{j+2}) \quad \text{and} \quad \hat{y}_i > 0;$$

notice that

$$\begin{aligned} X(t_i) - \hat{x}_i - \hat{y}_i &\overset{(A.6),(A.8)}{=} -A_2(t_i, t_{j+1}) - B_2(t_i, t_{j+1}) \\ &\overset{(A.7),(A.9)}{=} \frac{\Gamma}{2C(t_{i+1}, 0)} \\ &< \bar{X}(t_{i+1}, t_{i+2}). \end{aligned}$$

Hence, $(\hat{x}_i, \hat{y}_i)$ is the optimal strategy as long as

$$X(t_i) \leq \frac{\bar{X}(t_{i+1}, t_{j+1}) + A_2(t_i, t_{j+1})}{B_1(t_i, t_{j+1})} \overset{(A.7),(A.8),(A.11)}{=} \bar{X}(t_i, t_{j+2})$$

as desired. Plugging $(\hat{x}_i, \hat{y}_i)$ into $\hat{v}(X(t_i), \cdot)$, we obtain

$$v(t_i, X(t_i)) = C_1(t_i, t_{j+1}) X(t_i)^2 + C_2(t_i, t_{j+1}) X(t_i) + C_3(t_i, t_{j+1}),$$

which finishes the proof.

□

In the following we deduce properties of the optimal strategy and the functions $\bar{X}$ defined in Theorem 1.4.4. We discuss these properties in more detail in Section 1.4.5. Additionally, the properties are useful for finding closed form solutions to the coefficients $A_1, \ldots, C_3$ and the function $\bar{X}$ in Section 1.4.6.

Corollary 1.4.5. *Let $i = 0, \ldots, N-1$.*

(i) If $|X(t_i)| \leq \bar{X}(t_i, t_{i+1})$, then

$$\begin{aligned} y^*(t_i, X(t_i)) &= 0, \\ x^*(t_i, X(t_i)) &= A(t_i, 0) X(t_i), \\ v(t_i, X(t_i)) &= C(t_i, 0) X(t_i)^2 \end{aligned}$$

and

$$X(t_i) - x^*(t_i, X(t_i)) \leq \bar{X}(t_{i+1}, t_{i+2}).$$

(ii) If $|X(t_i)| \geq \bar{X}(t_i, t_N)$, then

$$\begin{aligned} A_1(t_i, t_N) &= A(t_i, p), \\ B_1(t_i, t_N) &= B(t_i, p), \\ C_1(t_i, t_N) &= C(t_i, p); \end{aligned}$$

recall that $A(t_i, p)$, $B(t_i, p)$ and $C(t_i, p)$ are the coefficients of the optimal strategy respectively the value function for optimal liquidation without adverse selection (cf. Equations (1.31) - (1.33)*).*

(iii) If $j = i + 1, \dots, N$ and $X(t_i) \in (\bar{X}(t_i, t_j), \bar{X}(t_i, t_{j+1}))$, then

$$x^*(t_i, X(t_i)), \quad y^*(t_i, X(t_i)) \quad \text{and} \quad X(t_i) - x^*(t_i, X(t_i))$$

are increasing in $X(t_i)$ and

$$X(t_i) - x^*(t_i, X(t_i)) - y^*(t_i, X(t_i)) = \frac{\Gamma}{2C(t_{i+1}, 0)} < \bar{X}(t_{i+1}, t_{i+2}),$$
$$X(t_i) - x^*(t_i, X(t_i)) \in (\bar{X}(t_{i+1}, t_j), \bar{X}(t_{i+1}, t_{j+1})).$$

(iv) For $j = i + 2, \dots, N$,

$$\bar{X}(t_i, t_j) = \min \{X \geq \bar{X}(t_i, t_{j-1}) \,|\, X - x^*(t_i, X) = \bar{X}(t_{i+1}, t_j)\}.$$

Proof. Assertion (i) follows directly from Theorem 1.4.4.

Note that the recursions for $C_1(\cdot, t_N)$, $A_1(\cdot, t_N)$ and $B_1(\cdot, t_N)$ given by Equations (A.3), (A.6) and (A.8) are the same as the recursions for $C(\cdot, p)$, $A(\cdot, p)$ and $B(\cdot, p)$ given by (1.34) and (1.35). The fact that

$$C(t_N, p) = C(t_N, t_N)$$

thus establishes (ii).

Due to the strict convexity of $v(t_k, \cdot)$, $k = i, \dots, N$ (cf. Theorem 1.4.2), we have

$$C_1(t_k, t_j) > 0, \quad k = i, \dots, N,$$

and thus by Equations (A.6) and (A.8) that

$$A_1(t_i, t_j) > 0 \quad \text{and} \quad B_1(t_i, t_j) > 0.$$

Therefore, $x^*(t_i, X(t_i))$ and $y^*(t_i, X(t_i))$ are strictly increasing in $X(t_i)$. Strict monotonicity of $X(t_i) - x^*(t_i, X(t_i))$ follows from the fact that

$$X(t_i) - x^*(t_i, X(t_i)) = B_1(t_i, t_j) X(t_i) - A_2(t_i, t_j).$$

Using Equation (A.6) - (A.9), we obtain

$$\begin{aligned} X(t_i) - x^*(t_i, X(t_i)) - y^*(t_i, X(t_i)) &= -A_2(t_i, t_j) - B_2(t_i, t_j) \\ &= \frac{\Gamma}{2C(t_{i+1}, 0)} \\ &< \frac{\Gamma(C(t_{i+2}, 0) + \Lambda)}{2C(t_{i+2}, 0)\Lambda} \end{aligned} \tag{1.72}$$

$$= \bar{X}(t_{i+1}, t_{i+2}),$$

where Inequality (1.72) follows from the fact that

$$C(t_{i+1}, 0) = \alpha\Sigma + \frac{\Lambda C(t_{i+2}, 0)}{\Lambda + C(t_{i+2}, 0)}$$

by Equation (1.35).

Finally, let $X \in (\bar{X}(t_i, t_{j-1}), \bar{X}(t_i, t_j))$ and note that $X - x^*(t_i, X)$ is strictly increasing in X. Then

$$X - x^*(t_i, X) = B_1(t_i, t_{j-1})X - A_2(t_i, t_{j-1}).$$

The assertion follows as

$$\lim_{X \to \bar{X}(t_i, t_j)-} X - x^*(t_i, X) = \bar{X}(t_{i+1}, t_j)$$

by Equations (A.7), (A.8) and (A.11). □

Risk-neutral investors

In Theorem 1.4.4 we assumed $\alpha\Sigma > 0$. Here, we treat the case of a risk-neutral investor

$$\alpha\Sigma = 0.$$

For trading without dark pools, it is optimal to liquidate the position evenly, and the optimal strategy and the value function are given by (cf. the Recursions (1.34) and (1.35))

$$\begin{aligned} C(t_i, 0)X(t_i)^2 &= \frac{\Lambda}{N + 1 - i} X(t_i)^2, \\ A(t_i, 0)X(t_i) &= \frac{1}{N + 1 - i} X(t_i). \end{aligned}$$

For obtaining the structure of the value function and the optimal strategy for risk-neutral trading with adverse selection, we let $\alpha\Sigma$ tend to zero and obtain that the boundary, below which the dark pool is not used, is given by

$$\bar{X}(t_i, t_{i+1}) = \frac{\Gamma(C(t_{i+1}, 0) + \Lambda)}{2C(t_{i+1}, 0)\Lambda} = \frac{\Gamma(N + 1 - i)}{2\Lambda}.$$

By continuity of the optimal strategy, we obtain

$$\bar{X}(t_i, t_{i+1}) - x^*(t_i, \bar{X}(t_i, t_{i+1})) = \bar{X}(t_{i+1}, t_{i+2}).$$

In contrast to the case of a risk-averse investor, the position does not *cross* the boundary. This suggests that for all asset positions

$$X(t_i) > \bar{X}(t_i, t_{i+1}),$$

the boundary is *never* crossed. Also, after a dark pool execution the position should be on the boundary:

$$X(t_i) - x^*(t_i, X(t_i)) - y^*(t_i, X(t_i)) = \frac{\Gamma}{2C(t_{i+1}, 0)} = \frac{\Gamma(N-i)}{2\Lambda} = \bar{X}(t_{i+1}, t_{i+2}).$$

The following result summarizes these findings and can be proven by the same line of reasoning as in the proof of Theorem 1.4.4 with appropriate modifications. Alternatively we can prove the result by taking the limit $\alpha\Sigma \to 0$ in the closed form solutions for $\bar{X}(t_i, t_j)$, $j = i+1, \ldots, N$, obtained in Section 1.4.6 below.

Proposition 1.4.6. *Let $\alpha\Sigma = 0$. For $i = 0, \ldots, N$ and asset position $X(t_i)$ at time t_i, the value function and the optimal strategy are given by*

$$v(t_i, X(t_i)) = \begin{cases} C(t_i, 0)X(t_i)^2 & \text{if } |X(t_i)| \le \bar{X}(t_i, t_{i+1}) \\ C_1(t_i, t_N)X(t_i)^2 + C_2(t_i, t_N)|X(t_i)| + C_3(t_i, t_N) & \text{else,} \end{cases}$$

$$x^*(t_i, X(t_i)) = \begin{cases} A(t_i, 0)X(t_i) & \text{if } |X(t_i)| \le \bar{X}(t_i, t_{i+1}) \\ A_1(t_i, t_N)X(t_i) + \operatorname{sgn}(X(t_i))A_2(t_i, t_N) & \text{else,} \end{cases}$$

$$y^*(t_i, X(t_i)) = \begin{cases} 0 & \text{if } |X(t_i)| \le \bar{X}(t_i, t_{i+1}) \\ B_1(t_i, t_N)X(t_i) + \operatorname{sgn}(X(t_i))B_2(t_i, t_N) & \text{else} \end{cases}$$

for $A_1, \ldots, C_1$ as in Appendix A.1. In particular,

$$\begin{aligned} X(t_i) - x^*(t_i, X(t_i)) &> \bar{X}(t_{i+1}, t_{i+2}), \\ X(t_i) - x^*(t_i, X(t_i)) - y^*(t_i, X(t_i)) &= \bar{X}(t_{i+1}, t_{i+2}) \end{aligned}$$

for $|X(t_i)| > \bar{X}(t_i, t_{i+1})$.

1.4.5. Properties of the optimal strategy

Theorem 1.4.4 and Corollary 1.4.5 provide the structure of the value function and the optimal strategy. In this section we discuss these properties and deduce additional properties.

For simplicity we assume $X(t_i) \ge 0$ and $\alpha\Sigma > 0$ throughout the section. By symmetry (cf. Theorem 1.4.4 (ii)) all properties transfer to negative asset positions $X(t_i) < 0$. It is always clear from the context whether corresponding properties also apply for risk-neutral investors.

General properties of the optimal strategy

Theorem 1.4.4 and Corollary 1.4.5 (i) stress the importance of the boundary $\bar{X}(t_i, t_{i+1})$ for the optimal strategy. The dark pool is only used at time t_i if the asset position $X(t_i)$ is above $\bar{X}(t_i, t_{i+1})$. If

$$X(t_i) < \bar{X}(t_i, t_{i+1}),$$

it is optimal to place no order in the dark pool and the optimal strategy in the primary venue is the same as the one without dark pool. This intuition is reflected by practitioners' rules of thumb which often do not use dark pools for orders smaller than a certain threshold. Even for asset positions

$$X(t_i) > \bar{X}(t_i, t_{i+1}),$$

it is no longer optimal to place the full remainder of the position in the dark pool (cf. Section 1.3.3) by Corollary 1.4.5 (iii). The trader wishes to keep a fraction of her position in order to not miss out a possible favorable price move completely. However, a dark pool execution at time t_j will cause the position to cross the boundary for the next trading interval $\bar{X}(t_{j+1}, t_{j+2})$, and consequently it is optimal to not use the dark pool afterwards (cf. Corollary 1.4.5 (iii) again). If

$$X(t_i) \in (\bar{X}(t_i, t_j), \bar{X}(t_i, t_{j+1}))$$

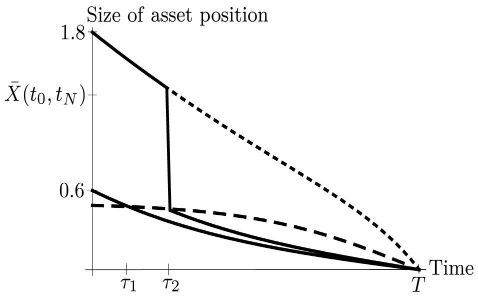
for some $j = i+1, \ldots, N$, the dark pool is only used at times $t_i, \ldots, t_{j-1}$ until execution. From time t_j respectively after execution, the dark pool is not used anymore as $X(t_j) < \bar{X}(t_j, t_{j+1})$ (cf. Corollary 1.4.5 (iii)). Only if

$$X(t_i) > \bar{X}(t_i, t_N)$$

and if orders in the dark pool are never executed, the optimal asset position stays above the boundary $\bar{X}(t_j, t_{j+1})$ throughout the whole time horizon $[0, T]$, and in this case it is optimal to use the dark pool at all trading times $t_i, \ldots, t_{N-1}$ until execution. Note also that by Corollary 1.4.5 (ii), adverse selection only has a second order impact on the optimal strategy and the value function for very large

$$X(t_i) >> \bar{X}(t_i, t_N).$$

We illustrate these properties by Figure 1.11. The left picture shows the optimal strategies for two initial asset positions. A larger one ($X(t_0) = 1.8$) which lies above $\bar{X}(t_0, t_N) = \bar{X}(t_0, t_{100}) = 1.27$ and a smaller one ($X(t_0) = 0.6$) which lies between $\bar{X}(t_0, t_{12})$ and $\bar{X}(t_0, t_{13})$. Consequently, the larger asset position crosses the boundary $\bar{X}(t_i, t_{i+1})$ (dashed line) only if the order in the dark pool is executed (which happens at time τ_2 in the displayed scenario). The smaller one crosses the boundary after the twelfth trading period (at time τ_1) if no order in the dark pool is executed before. Compared to the optimal strategies if no adverse selection is expected (right picture, solid lines), the trading speed in the primary venue is initially faster but still slower than the trading speed of the optimal strategy without dark pool (thin lines). Additionally, the order in the dark pool is smaller than the remainder of the position, and after execution in the dark pool the investor trades out of the rest solely in the primary venue. The reason for both properties is the fact that trading in the dark pool is not entirely free anymore and thus (in relation to it) trading is cheaper in the primary venue.



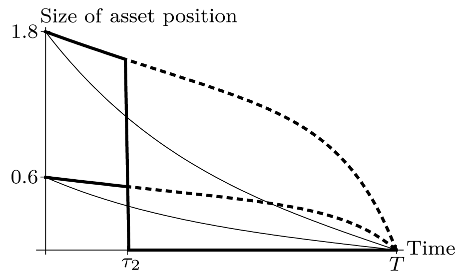


Figure 1.11.: The left picture shows the optimal strategies for a smaller and a larger asset position if adverse selection is expected. The right picture shows the respective strategies if adverse selection is not expected. In all cases we consider the scenario where the order in the dark pool is executed at time τ_2. The solid lines denote the optimal strategies for this scenario, the dotted those for the scenario where orders in the dark pool are never executed. The dashed line in the left picture reflects the boundary $\bar{X}(t_i, t_{i+1})$ below which the optimal order in the dark pool is zero. The thin lines in the right pictures denote the optimal strategies without dark pools. $N = 100$, $\Lambda = 100$, $\Sigma = \frac{1}{100}$, $p = \frac{6}{100}$, $\alpha = 4$, $\Gamma = 2$.

Coefficients of the value function and the optimal strategy

As an immediate consequence of Theorem 1.4.4, we obtain the following corollary. It is concerned with the coefficients of the value function and the optimal strategy.

Corollary 1.4.7. *Let $i = 0, \dots N$, $j = i, ..., N$.*

(i)

$$C_1(t_i, t_j), C_2(t_i, t_j) \geq 0 \quad \text{and} \quad C_3(t_i, t_j) \leq 0.$$

$C_1(t_i, \cdot)$ and $C_3(t_i, \cdot)$ are decreasing, and $C_2(t_i, \cdot)$ is increasing.

(ii)

$$A_1(t_i, t_j), A_2(t_i, t_j), B_1(t_i, t_j) \geq 0 \quad \text{and} \quad B_2(t_i, t_j) \leq 0.$$

$A_1(t_i, \cdot)$ and $B_2(t_i, \cdot)$ are decreasing, and $A_2(t_i, \cdot)$ and $B_1(t_i, \cdot)$ are increasing. In particular, $x^*(t_i, \cdot)$ is concave and $y^*(t_i, \cdot)$ is convex.

Proof. Both assertions are straightforward backward inductions using the respective recursions from Appendix A.1. □

Figure 1.12 illustrates the dependence of the optimal order size in the primary venue (solid line) and in the dark pool (dashed line) on the asset position. As shown in Corollary 1.4.7, the order size in the primary venue is concave in $X(t_0)$ and the order size in the dark pool is convex in $X(t_0)$.

Dependence on Γ

In Section 1.3.3 we showed that a high probability of execution in the dark pool slows down trading in the primary venue initially. Intuitively, adverse selection should have the

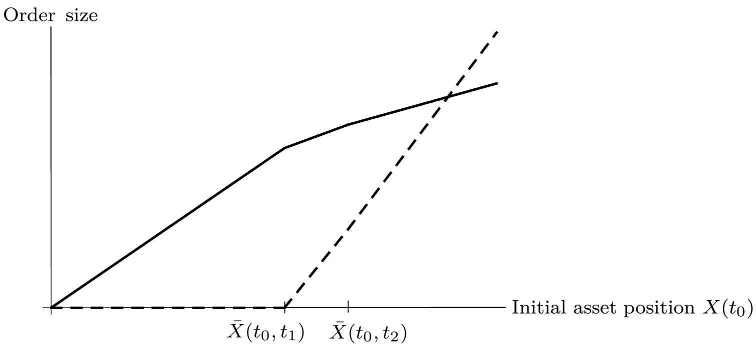


Figure 1.12.: Optimal order sizes in the dark pool and the primary venue at time t_0 depending on the asset position $X(t_0)$. The solid line denotes the optimal order size in the primary venue $x^*(t_0, X(t_0))$, the dashed line the optimal order size in the dark pool $y^*(t_0, X(t_0))$. $N = 2$, $\Lambda = 10$, $\Sigma = \frac{1}{2}$, $p = 0.6$, $\alpha = 4$, $\Gamma = 10$.

opposite effect: higher adverse selection should speed up trading in the primary venue as trading in the dark pool is relatively more expensive and thus waiting for execution is less attractive (cf. also Figure 1.11 and the corresponding discussion). The following proposition confirms that this intuition is correct.

In order to stress the dependence of the value function and the optimal strategy on the parameter Γ, we first introduce the following notation: for a setting with adverse selection $\Gamma \geq 0$, we denote the optimal strategy at time t_i by

$$(x^*(t_i, X(t_i), \Gamma), y^*(t_i, X(t_i), \Gamma)).$$

In a similar fashion we characterize optimal trajectories, value functions etc. Similarly as in the proof of Lemma 1.4.9, we define $X^{\text{ne}}(t_j, \Gamma)$ $(j = i, \ldots, N)$ recursively by

$$X^{\text{ne}}(t_i, \Gamma) = X(t_i) \quad \text{and} \quad X^{\text{ne}}(t_j, \Gamma) = X^{\text{ne}}(t_{j-1}, \Gamma) - x^*(t_j, X^{\text{ne}}(t_{j-1}, \Gamma), \Gamma) \quad \text{for } j > i.$$

Proposition 1.4.8. *Let $i = 0, \ldots, N-1$ and $X(t_i) > 0$ as before. We define*

$$\Gamma_0 := \frac{2C(t_{i+1}, 0)\Lambda}{C(t_{i+1}, 0) + \Lambda} X(t_i).$$

(i) For $j = i+1, \ldots, N$, $\bar{X}(t_i, t_j, \cdot)$ is strictly increasing for $\Gamma \geq 0$.

(ii) $v(t_i, X(t_i), \cdot)$ is strictly increasing for $\Gamma < \Gamma_0$ and constant for $\Gamma \geq \Gamma_0$.

(iii) $x^*(t_i, X(t_i), \cdot)$ is strictly increasing for $\Gamma < \Gamma_0$ and constant for $\Gamma \geq \Gamma_0$.

(iv) $y^*(t_i, X(t_i), \cdot)$ is strictly decreasing for $\Gamma < \Gamma_0$ and constant for $\Gamma \geq \Gamma_0$.

(v) For $j = i+1, \ldots, N$, $X^{\text{ne}}(t_j, \cdot)$ is strictly decreasing for $\Gamma < \Gamma_0$ and constant for $\Gamma \geq \Gamma_0$.

Proof. (i) Note first that C_1 is independent of Γ by Equations (A.1) and (A.3). By backward induction on i using Equation (A.4), we deduce that for $i = 0, \dots, N-1$ and $j = i+1, \dots, N$, $C_2(t_i, t_j, \cdot)$ is strictly increasing in Γ. The assertion follows by another backward induction on i using Equation (A.11).

(ii) For $\Gamma < \Gamma_0$, adverse selection is small enough to ensure that

$$X(t_i) > \frac{\Gamma(C(t_{i+1}, 0) + \Lambda)}{2C(t_{i+1}, 0)\Lambda}$$

and thus

$$y^*(t_i, X(t_i), \Gamma) > 0.$$

Applying the strategy $(x^*(t_i, X(t_i), \Gamma), y^*(t_i, X(t_i), \Gamma))$ to a setting with adverse selection $\tilde{\Gamma} < \Gamma$ generates hence strictly less costs. This implies by backward induction that

$$v(t_i, X(t_i), \tilde{\Gamma}) < v(t_i, X(t_i), \Gamma).$$

(iii) Note first that A_1 is independent of Γ by Equation (A.1) and (A.6) (recall that C_1 is independent of Γ). As $C_2(t_i, t_j, \cdot)$ is strictly increasing in Γ (cf. the proof of (i)), $A_2(t_i, t_j)$ is strictly increasing in Γ for $i = 0, \dots, N-1$, $j = i+1, \dots, N$ by Equation (A.7). Let now $\tilde{\Gamma} < \Gamma < \Gamma_0$ and assume that

$$X(t_i) \in [\bar{X}(t_i, t_j, \Gamma), \bar{X}(t_i, t_{j+1}, \Gamma)) \quad \text{for} \quad j \in \{i+1, \dots, N\}.$$

By (i), we have $X(t_i) \in [\bar{X}(t_i, t_h, \tilde{\Gamma}), \bar{X}(t_i, t_{h+1}), \tilde{\Gamma})$ for $h \geq j$. Therefore (and by the monotonicity of A_2),

$$\begin{aligned} x^*(t_i, X(t_i), \tilde{\Gamma}) &= A_1(t_i, t_h, \tilde{\Gamma})X(t_i) + A_2(t_i, t_h, \tilde{\Gamma}) \\ &\leq A_1(t_i, t_j, \tilde{\Gamma})X(t_i) + A_2(t_i, t_j, \tilde{\Gamma}) \qquad (1.73) \\ &< A_1(t_i, t_j, \Gamma)X(t_i) + A_2(t_i, t_j, \Gamma) \\ &= x^*(t_i, X(t_i), \Gamma), \end{aligned}$$

where Inequality (1.73) follows from the concavity of $x^*(t_i, \cdot, \tilde{\Gamma})$ (cf. Corollary 1.4.7 (ii)).

(iv) By Corollary 1.4.5 (iii),

$$x^*(t_i, X(t_i), \cdot) + y^*(t_i, X(t_i), \cdot)$$

is strictly decreasing in Γ, and hence $y^*(t_i, X(t_i), \Gamma)$ is strictly decreasing in Γ by (iii).

(v) Let $\tilde{\Gamma} < \Gamma < \Gamma_0$. By forward induction on j, we deduce from (iii) that for $j > i$,

$$X^{\text{ne}}(t_j, \Gamma) = X^{\text{ne}}(t_{i-1}, \Gamma) - x^*(t_{i-1}, X^{\text{ne}}(t_{i-1}, \Gamma), \Gamma)$$

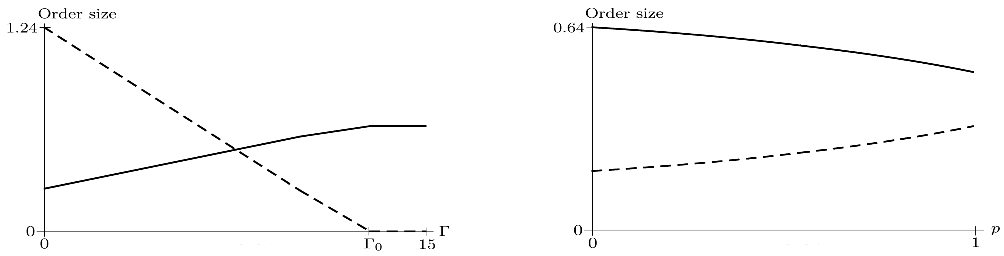


Figure 1.13.: Dependence of the optimal order size in the primary venue (solid lines) and in the dark pool (dashed lines) on adverse selection (left picture) and on the probability of execution in the dark pool (right picture), respectively. $X(t_0) = 1.5$, $N = 2$, $\Lambda = 10$, $\Sigma = \frac{1}{2}$, $\alpha = 4$; $p = 0.6$ in the left picture, $\Gamma = 10$ in the right picture.

$$\begin{aligned}
&\leq X^{\text{ne}}(t_{i-1}, \tilde{\Gamma}) - x^*(t_i, X^{\text{ne}}(t_{i-1}, \tilde{\Gamma}), \Gamma) \qquad (1.74)\\
&< X^{\text{ne}}(t_{i-1}, \tilde{\Gamma}) - x^*(t_i, X^{\text{ne}}(t_{i-1}, \tilde{\Gamma}), \tilde{\Gamma})\\
&= X^{\text{ne}}(t_i, \tilde{\Gamma}),
\end{aligned}$$

where (1.74) follows from the induction hypothesis and the fact that $X(t_i) - x^*(t_i, X(t_i))$ is strictly increasing in $X(t_i)$ (cf. Corollary 1.4.5 (iii)).

□

Figure 1.13 illustrates the dependence of the trading speed in the primary venue on adverse selection (left picture) and on the probability of execution (right picture). As we have shown in Proposition 1.4.8, large adverse selection speeds up trading in the primary venue and decreases the optimal order size in the dark pool. On the other hand, high execution probability in the dark pools slows down trading in the primary venue and increases the optimal order size in the dark pool.

1.4.6. Closed form solutions

With the results about the optimal strategy given in Theorem 1.4.4 and Corollary 1.4.5, we are able to derive inhomogeneous linear difference equations for the functions $\bar{X}(\cdot, t_j)$, which we solve explicitly in Corollary 1.4.10. We can deduce closed form solutions for the coefficients $A_1, \ldots, C_3$.

To this end, we first require the following lemma which reduces the set of possibly optimal strategies.

Lemma 1.4.9. *Let $j = 1, \ldots, N$ and $X(t_0) = \bar{X}(t_0, t_j)$. For an admissible liquidation strategy $(\tilde{x}, \tilde{y}) \in \mathbb{A}(t_0, X(t_0))$, we define the trading trajectory for the scenario, where no dark pool orders are executed by*

$$\tilde{X}(t_i), \quad i = 0, \ldots, N.$$

Let $(\tilde{x}, \tilde{y})$ be optimal. If no dark pool order is executed before t_{j-2}, then

(i)
$$\tilde{X}(t_{j-1}) = \bar{X}(t_{j-1}, t_j) = \frac{\Gamma(C(t_j, 0) + \Lambda)}{2C(t_j, 0)\Lambda},$$

(ii)
$$\tilde{X}(t_j) = \frac{\Gamma}{2C(t_j, 0)},$$

(iii)
$$\tilde{X}(t_i) - \tilde{x}(t_i) - \tilde{y}(t_i) = \frac{\Gamma}{2C(t_{j+1}, 0)}$$

for $i = 0, \ldots, j-2$*, and*

(iv)
$$\tilde{x}(t_i) = A(t_i, 0)\tilde{X}(t_i), \quad \tilde{y}(t_i) = 0$$

for $i \geq j$*.*

If a dark pool order is executed at time t_k*,* $k = 0, \ldots, j-2$*, then*

(v) the optimal strategy without dark pool is applied at trading times $t_{k+1}, \ldots, t_N$*.*

Proof. Let $X^{\mathrm{ne}}(t_i)$ be defined recursively by

$$X^{\mathrm{ne}}(t_0) = X(t_0) \quad \text{and} \quad X^{\mathrm{ne}}(t_i) = X^{\mathrm{ne}}(t_{i-1}) - x^*(t_i, X^{\mathrm{ne}}(t_{i-1})) \quad \text{for } i > 0. \tag{1.75}$$

By Theorem 1.4.4, the optimal strategy $(\tilde{x}, \tilde{y})$ must satisfy

$$\tilde{X}(t_i) = X^{\mathrm{ne}}(t_i)$$

for all $i = 0, \ldots, N$. We have

$$X^{\mathrm{ne}}(t_i) = \bar{X}(t_i, t_j)$$

for $i = 0, \ldots, j-1$ by Corollary 1.4.5 (iv), in particular

$$X^{\mathrm{ne}}(t_{j-1}) = \bar{X}(t_{j-1}, t_j) = \frac{\Gamma(C(t_j, 0) + \Lambda)}{2C(t_j, 0)\Lambda},$$

and (i) follows. Assertion (ii) follows as

$$\begin{aligned} X^{\mathrm{ne}}(t_j) &= X^{\mathrm{ne}}(t_{j-1}) - x^*(t_{j-1}, X^{\mathrm{ne}}(t_{j-1})) = (1 - A(t_{j-1}, 0))X^{\mathrm{ne}}(t_{j-1}) \\ &= \frac{\Gamma}{2C(t_j, 0)}. \end{aligned}$$

Assertion (iii) follows from Corollary 1.4.5 (iii), Assertions (iv) and (v) follow directly from Corollary 1.4.5 (i) and (iii). $\square$

We denote the set of admissible strategies satisfying (i) - (v) by $\bar{\mathbb{A}}(t_0, X(t_0))$*.*

We are now able to derive explicit formulae for $\bar{X}(t_i, t_j)$ ($i = 0, \ldots, N-1$, $j = i, \ldots, N$) by minimizing the costs over those strategies only that fulfill (i) - (v).

Corollary 1.4.10. *Let $i = 0, \ldots, N-1$, $j = i, \ldots, N$. Then*

$$\begin{aligned}\bar{X}(t_i, t_j) = {} & \frac{1}{2\alpha\Sigma(1-p)\sinh(\kappa(p))} \\ & \cdot \Big(2\alpha\Sigma\Big(d(t_j)\sinh((j-i)\kappa(p))\sqrt{1-p}^{\,j-i+1} \\ & \qquad - e(t_j)\sinh((j-i-1)\kappa(p))\sqrt{1-p}^{\,j-i+2}\Big) \\ & \quad + \Gamma p\Big(\sqrt{1-p}^{\,j-i-1}\sinh((j-i)\kappa(p)) \\ & \qquad - \sqrt{1-p}^{\,j-i}\sinh((j-i-1)\kappa(p)) - \sinh(\kappa(p))\Big)\Big),\end{aligned} \tag{1.76}$$

where

$$d(t_j) := \frac{\Gamma(C(t_j, 0) + \Lambda)}{2C(t_j, 0)\Lambda}, \quad e(t_j) := \frac{\Gamma}{2C(t_j, 0)}$$

for $j < N$ and

$$d(t_N) := e(t_N) := 0.$$

Closed form solutions for the coefficients $A_1, \ldots, C_3$ are given in Appendix A.2 (Equations (A.12) - (A.24)*).*

Proof. Let $j = 2, \ldots, N$ and $X(t_0) = \bar{X}(t_0, t_j)$. Define $X^{\mathrm{ne}}(t_i)$ as in (1.75) and consider the cost functional $J(t_0, X(t_0), \cdot) : \mathbb{A}(t_0, X(t_0)) \longrightarrow \mathbb{R}$ given by

$$J(t_0, X(t_0), (x, y)) := \mathbb{E}\Big[\sum_{j=0}^{N} \Lambda x(t_j)^2 + p\Gamma|y(t_j)|\Big] + \alpha\mathbb{E}\Big[\sum_{j=0}^{N} \Sigma X(t_j)^2\Big].$$

By Lemma 1.4.9, we can replace $\mathbb{A}$ by $\bar{\mathbb{A}}$:

$$v(t_0, X(t_0)) = \inf_{(x,y)\in\mathbb{A}(t_0,X(t_0))} J(t_0, X(t_0), (x, y)) = \inf_{(x,y)\in\bar{\mathbb{A}}(t_0,X(t_0))} J(t_0, X(t_0), (x, y)).$$

Let now $(\tilde{x}, \tilde{y}) \in \bar{\mathbb{A}}(t_0, X(t_0))$ and define $\tilde{X}(t_i)$ as in Lemma 1.4.9. By the definition of $\bar{\mathbb{A}}$, we have

$$\tilde{X}(t_0) = X(t_0), \quad \tilde{X}(t_{j-1}) = \frac{\Gamma(C(t_j, 0) + \Lambda)}{2C(t_j, 0)\Lambda}, \quad \tilde{X}(t_j) = \frac{\Gamma}{2C(t_{j+1}, 0)}.$$

The costs of the strategy $(\tilde{x}, \tilde{y})$ are given by

$$\begin{aligned}& J(t_0, X(t_0), (\tilde{x}, \tilde{y})) \\ & \quad = \sum_{i=0}^{j-2} (1-p)^i \Big(\Lambda(\tilde{X}(t_i) - \tilde{X}(t_{i+1}))^2 + \alpha\Sigma\tilde{X}(t_i)^2 + p\Gamma\Big(\tilde{X}(t_{i+1}) - \tfrac{\Gamma}{2C(t_{i+1},0)}\Big)\Big)\end{aligned}$$

$$
\begin{aligned}
&+\sum_{i=0}^{j-2}(1-p)^i pC(t_{i+1},0)\Big(\tfrac{\Gamma}{2C(t_{i+1},0)}\Big)^2 \\
&+(1-p)^{j-1}\Big(\Lambda(\tilde{X}(t_{j-1})-\tilde{X}(t_j))^2+\alpha\Sigma\tilde{X}(t_{j-1})^2+C(t_j,0)\Big(\tfrac{\Gamma}{2C(t_j,0)}\Big)^2\Big) \quad (1.77)\\
&=: U(\tilde{X}(t_1),\dots,\tilde{X}(t_{j-2})).
\end{aligned}
$$

Let us briefly comment on the terms in Equation (1.77).

(i) The terms

$$
\Lambda(\tilde{X}(t_i)-\tilde{X}(t_{i+1}))^2, \quad \alpha\Sigma(\tilde{X}(t_i))^2, \quad p\Gamma\Big(\tilde{X}(t_{i+1})-\frac{\Gamma}{2C(t_{i+1},0)}\Big), \quad i=0,\dots,j-2,
$$

reflect the impact, risk respectively adverse selection costs at time t_i in the scenarios where no dark pool order is executed before time t_i (cf. Lemma 1.4.9 (iii) for the adverse selection costs). This happens with probability $(1-p)^i$.

(ii) The term

$$
C(t_{i+1},0)\Big(\tfrac{\Gamma}{2C(t_{i+1},0)}\Big)^2, \quad i=0,\dots,j-2,
$$

reflects the costs for the remainder of the trading horizon (times $t_{i+1},\dots,t_N$) in the scenarios, where the dark pool order is executed at time t_i; the corresponding scenarios have probability $p(1-p)^i$.

(iii) The terms

$$
\Lambda(\tilde{X}(t_{j-1})-\tilde{X}(t_j))^2+\alpha\Sigma\tilde{X}(t_{j-1})^2, \quad C(t_j,0)\Big(\tfrac{\Gamma}{2C(t_j,0)}\Big)^2
$$

reflect the costs at time t_{j-1} respectively for the remainder of the trading horizon (times $t_j,\dots,t_N$) in the scenarios, where no dark pool order is executed until time t_{j-1}. This happens with probability $(1-p)^{j-1}$.

The optimal strategy (x^*,y^*) minimizes $J(t_0,X_0,\cdot)$ uniquely. Therefore,

$$
(X^{\mathrm{ne}}(t_1),\dots,X^{\mathrm{ne}}(t_{j-2}))
$$

minimizes U uniquely and solves the system of linear equations

$$
\frac{\partial U}{\partial x_i}(\tilde{X}(t_1),\dots,\tilde{X}(t_{j-2}))=0, \quad i=1,\dots,j-2.
$$

This is equivalent to the inhomogeneous linear difference equation

$$
\tilde{X}(t_i)\left(1+\frac{1}{1-p}+\frac{\alpha\Sigma}{\Lambda}\right)+\frac{\Gamma p}{2(1-p)\Lambda}=\tilde{X}(t_{i+1})+\frac{1}{1-p}\tilde{X}(t_{i-1}). \quad (1.78)
$$

By standard methods we compute a special solution of Equation (1.78):

$$X^{(1)}(t_i) = \frac{-\Gamma p}{2(1-p)\alpha\Sigma}$$

and two linearly independent solutions of the corresponding homogeneous linear difference equation:

$$X^{(2)}(t_i) = \frac{\exp(\kappa(p)(j-1-i))}{\sqrt{1-p}^{i}}, \quad X^{(3)}(t_i) = \frac{\exp(-\kappa(p)(j-1-i))}{\sqrt{1-p}^{i}}.$$

Consequently, the solutions of Equation (1.78) are given by

$$X^{(1)}(t_i) + aX^{(2)}(t_i) + bX^{(3)}(t_i)$$

for $a, b \in \mathbb{R}$.

$X^{\mathrm{ne}}(t_i) = \bar{X}(t_i, t_j)$ is the unique solution of Equation (1.78) satisfying the boundary conditions

$$\tilde{X}(t_{j-1}) = \frac{\Gamma(C(t_j, 0) + \Lambda)}{2C(t_j, 0)\Lambda}, \quad \tilde{X}(t_j) = \frac{\Gamma}{2C(t_j, 0)}.$$

Elementary but tedious algebraic manipulations confirm Equation (1.76).

We deduce closed form solutions for the coefficients $A_1, \ldots, C_3$ as follows. By Corollary 1.4.5 (iv), we have

$$\bar{X}(t_i, t_j) - A_1(t_i, t_j)\bar{X}(t_i, t_j) - A_2(t_i, t_j) = \bar{X}(t_{i+1}, t_j),$$
$$\bar{X}(t_i, t_{j+1}) - A_1(t_i, t_j)\bar{X}(t_i, t_{j+1}) - A_2(t_i, t_j) = \bar{X}(t_{i+1}, t_{j+1})$$

for $i = 0, \ldots, N-2$, $j = i+1, \ldots, N-1$. Here, we defined by abuse of notation

$$\bar{X}(t_{i+1}, t_{i+1}) := \frac{\Gamma}{2C(t_{i+2}, 0)} \neq 0.$$

Solving this system of linear equations in $A_1(t_i, t_j)$ and $A_2(t_i, t_j)$ yields Equations (A.12) and (A.13). Equations (A.14) - (A.17) follow from that directly with the corresponding recursions from Appendix A.1. Equations (A.18), (A.20) and (A.22) follow from Corollary 1.4.5 (ii). Therefore, Equation (A.19) follows as

$$\underbrace{\bar{X}(t_i, t_N) - A_1(t_i, t_N)\bar{X}(t_i, t_N)}_{=B_1(t_i, t_N)\bar{X}(t_i, t_N)} - A_2(t_i, t_N) = \bar{X}(t_{i+1}, t_N).$$

Equation (A.21) follows from Corollary 1.4.5 (iv) (recall that $A(t_i, p) + B(t_i, p) = 1$), and Equation (A.23) follows from Equation (A.7). Finally, Equation (A.24) follows by a backward induction on $i \leq N-1$ from Equations (A.1) and (A.5). $\square$

1.5. Trading prices in the dark pool

So far we have assumed that trades in the dark pool are executed at the unaffected price $\tilde{P}$. Within this section we assume instead that *dark pool orders are executed at the exchange quoted price* P, which includes the temporary market impact of the orders $x(t_i)$. As indicated in Section 1.1.2, this might be a more appropriate assumption for some dark pools. This results in profitable market manipulating strategies unless the model parameters are chosen with great care as we shall show in this section. For simplicity, we assume the *single asset model* described in Section 1.3.3 and furthermore assume that the investor is *risk-neutral* ($\alpha = 0$) in this section.

Market manipulation is a concern in all market models where a large trader's orders have a feedback effect on the execution price of her own orders. Huberman and Stanzl [2004] and Gatheral [2010] derive necessary conditions for market models that exclude profitable market manipulation at a primary exchange. Both papers disregard trading opportunities in dark pools. For the primary exchange, the market model introduced in Section 1.3 fulfills the requirements established in these papers, i.e., it is not possible to generate profits from market manipulation by trading *only at the primary exchange.* However, it might be possible to generate profits from market manipulation if orders are placed cleverly in parallel in the dark pool. It is unclear whether such profitable market manipulation strategies exist in reality; given that such strategies were used and had to be forbidden (see Gatheral [2010] for an exposition), such opportunities seem to be available at least sometimes. Nevertheless, we agree with Huberman and Stanzl [2004], Gatheral [2010], Alfonsi and Schied [2010] and Alfonsi et al. [2009] that an appropriate mathematical market model should exclude profitable market manipulation.

For the purposes of this section, we define market manipulation strategies in the following way.

Definition 1.5.1. *Let* $i = 0, \dots, N$ *and* $X(t_i) \in \mathbb{R}$. *We call a strategy* $(x(t_i), y(t_i))$ *a* market manipulation strategy *if*

$$\operatorname{sgn}(X(t_i)) \neq \operatorname{sgn}(x(t_i)) \quad \textit{or} \quad \operatorname{sgn}(X(t_i)) \neq \operatorname{sgn}(y(t_i)).$$

As we saw in Section 1.3.4, such orders can be attractive as risk mitigation tools in a multi asset setting. In the single asset setting of this section this justification does not apply, and we saw in Section 1.3.3 that if trades are executed in the dark pool at fundamental prices, then market manipulation as defined here is never optimal.

In the following we consider in particular a market manipulation strategy similar to the classical 'pump and dump' strategy[2]. In our market model, selling the stock at the primary exchange after artificially elevating its price ('pumping') cannot generate profits due to the associated price reaction. A liquidation in the dark pool however does not face such a price penalty. Consider the following strategy:

[2]"'Pump and dump' schemes, also known as 'hype and dump manipulation', involve the touting of a company's stock [...] . After pumping the stock, fraudsters make huge profits by selling their cheap stock into the market." (From http://www.sec.gov/answers/pumpdump.htm)

Assume that the initial asset position is zero and that the number of trading time points $N+1$ is divisible by four. From t_0 until $t_{(N+1)/4}$ the investor buys a stock quantity X at each point in time at the primary exchange. Simultaneously, she seeks to dump shares by placing a sell order for $\frac{(N+1)X}{2}$ in the dark pool until the order gets executed (if at all). At time $t_{(N+1)/4}$ the investor either holds a long or short position of $\frac{(N+1)X}{4}$ in the asset, which she liquidates at a constant rate over the remaining time points $t_{(N+1)/4}, \ldots, t_N$. The expected trading proceeds are then

$$
\begin{aligned}
&\mathbb{E}\Big[\sum_{i=0}^{N}(x(t_i)+z(t_i))P(t_i)\Big] \\
&\quad = \underbrace{\mathbb{E}\Big[\sum_{i=0}^{N}(x(t_i)+z(t_i))\tilde{P}(t_i)\Big]}_{=0} - \mathbb{E}\Big[\sum_{i=0}^{N}(x(t_i)+z(t_i))\Lambda x(t_i)\Big] \\
&\quad = \Lambda\Big(- (N+1)(\frac{1}{4}+\frac{3}{4}\frac{1}{9})X^2 + (1-(1-p)^{(N+1)/4})\frac{(N+1)X^2}{2}\Big) \\
&\quad = \Lambda(N+1)\Big(\frac{1}{6} - \frac{(1-p)^{(N+1)/4}}{2}\Big)X^2.
\end{aligned}
$$

The last expression is positive if the number of trading time points $N+1$ is large enough. Furthermore, the expected proceeds grow in the position sizing factor X: the larger the bets, the larger the expected proceeds. The following proposition summarizes the issues we found.

Proposition 1.5.2. *Assume that trades in the dark pool are executed at the market price P. If*

$$N+1 \geq \lfloor 4\log(1/3)/\log(1-p)\rfloor + 1,$$

then profitable market manipulation strategies exist and optimal strategies do not exist.

In Section 1.3.3 we assumed both infinite liquidity in the dark pool if trading is possible ($a(t_i), b(t_i) \in \{0, \infty\}$) and no adverse selection ($\epsilon(t_{i+1})$ independent of $a(t_i), b(t_i)$). We replace Assumption 1.3.1 (iii) by the following assumption.

Assumption 1.5.3. *Let $i = 0, \ldots, N-1$.*

(i) Liquidity in the dark pool is bounded:

$$a(t_i), b(t_i) \in \{0, L\}$$

for some $L \in (0, \infty)$.

(ii) There might be adverse selection:

$$\mathbb{E}[\epsilon(t_{i+1})|a(t_i) = L] = -\Gamma, \quad \mathbb{E}[\epsilon(t_{i+1})|b(t_i) = L] = \Gamma$$

with $\Gamma > 0$.

By limiting dark pool liquidity, market manipulating strategies with very large trades cannot be profitable. On the other hand, adverse selection makes market manipulation by small trades unprofitable. The following proposition shows that if Assumption 1.3.1 (iii) is replaced by Assumption 1.5.3 and adverse selection is sufficiently large, then the undesirable properties outlined in Proposition 1.5.2 disappear.

Proposition 1.5.4. *Let $i = 0, \dots, N$. Assume that trades in the dark pool are executed at the market price $P(t_j)$, $j = i, \dots, N$. We consider the following optimization problem:*

$$\bar{v}(t_i, X(t_i)) := \inf_{(x,y)\in\mathbb{A}(t_i,X(t_i))} \mathbb{E}\Big[\sum_{j=i}^{N} \big(x(t_j) + z(t_j)\big)^\top \big(\tilde{P}(t_i) - P(t_j)\big)\Big]. \tag{1.79}$$

If

$$\Gamma > \Lambda L, \tag{1.80}$$

then there exist optimal strategies realizing the minimum in Equation (1.79) *and these are not market manipulating.*

Proof. We prove existence of an optimal trading strategy together with symmetry,

$$\bar{v}(t_i, -X(t_i)) = \bar{v}(t_i, X(t_i)),$$

monotonicity,

$$|X(t_i)| \le |Y(t_i)| \Rightarrow \bar{v}(t_i, X(t_i)) \le \bar{v}(t_i, Y(t_i))$$

and continuity of the value function $\bar{v}(t_i, \cdot)$ by backward induction (cf. Theorem 1.4.2 and its proof). The case $i = N$ is clear. For $i < N$, Equation (1.79) yields the Bellman equation

$$\begin{aligned}\bar{v}(t_i, X(t_i)) = &\inf_{(x,y)\in\mathbb{R}\times\mathbb{R},\,|y|\le L} \big\{\Lambda x^2 + p|y|\Gamma + p\Lambda xy + p\bar{v}(t_{i+1}, X(t_i) - x - y) \\ &\qquad + (1-p)\bar{v}(t_{i+1}, X(t_i) - x)\big\}. \\ =: &\inf_{(x,y)\in\mathbb{R}\times\mathbb{R},\,|y|\le L} \tilde{v}(X(t_i), x, y).\end{aligned} \tag{1.81}$$

Symmetry follows immediately. For monotonicity, we can therefore assume without loss of generality that $0 \le X(t_i) \le Y(t_i)$. Let $(x(t_i), y(t_i)) \in \mathbb{R}^2$ be an admissible pair of orders. Let us first assume that

$$\operatorname{sgn}(x(t_i)) = \operatorname{sgn}(y(t_i)) \quad \text{or} \quad x(t_i)\cdot y(t_i) = 0.$$

As in the proof of Theorem 1.4.2, we define:

$$\tilde{x}(t_i) := \begin{cases} x(t_i) & \text{if } X(t_i) - x(t_i) \ge 0 \\ X(t_i) & \text{else,} \end{cases}$$

$$\tilde{y}(t_i) := \begin{cases} y(t_i) & \text{if } X(t_i) - x(t_i), X(t_i) - x(t_i) - y(t_i) \geq 0 \\ X(t_i) - x(t_i) & \text{if } X(t_i) - x(t_i) \geq 0,\ X(t_i) - x(t_i) - y(t_i) < 0 \\ 0 & \text{else.} \end{cases}$$

In all three possible cases a simple argument using the induction hypothesis establishes

$$\tilde{v}(X(t_i), \tilde{x}(t_i), \tilde{y}(t_i)) \leq \tilde{v}(Y(t_i), x(t_i), y(t_i)).$$

Let therefore

$$\operatorname{sgn}(x(t_i)) \neq \operatorname{sgn}(y(t_i)).$$

Assume that

$$\bar{v}(t_{i+1}, Y(t_i) - x(t_i)) < \bar{v}(t_{i+1}, Y(t_i) - x(t_i) - y(t_i)) + \Lambda x(t_i) y(t_i) + \Gamma |y(t_i)|.$$

Then orders

$$\bar{x}(t_i) = x(t_i) \quad \text{and} \quad \bar{y}(t_i) = 0$$

result in lower costs:

$$\Lambda x(t_i)^2 + \bar{v}(t_{i+1}, Y(t_i) - x(t_i)),$$

and therefore we obtain as before

$$\tilde{v}(X(t_i), \tilde{\bar{x}}(t_i), \tilde{\bar{y}}(t_i)) \leq \tilde{v}(Y(t_i), \bar{x}(t_i), \bar{y}(t_i)) < \tilde{v}(Y(t_i), x(t_i), y(t_i)).$$

Otherwise,

$$\bar{v}(t_{i+1}, X(t_i) - x(t_i)) \geq \bar{v}(t_{i+1}, X(t_i) - x(t_i) - y(t_i)) + \Lambda x(t_i) y(t_i) + \Gamma |y(t_i)|$$

and orders

$$\bar{x}(t_i) = x(t_i) + y(t_i) \quad \text{and} \quad \bar{y}(t_i) = 0$$

result in lower costs:

$$\begin{aligned} &\Lambda x(t_i)^2 + \Lambda y(t_i)^2 + 2\Lambda x(t_i) y(t_i) + \bar{v}(t_{i+1}, X(t_i) - x(t_i) - y(t_i)) \\ &\qquad < \Lambda x(t_i)^2 + \Lambda y(t_i) x(t_i) + \Gamma |y(t_i)| + \bar{v}(t_{i+1}, X(t_i) - x(t_i) - y(t_i)), \end{aligned} \tag{1.82}$$

where Inequality (1.82) follows from Condition (1.80), $\operatorname{sgn}(x^*(t_i)) \neq \operatorname{sgn}(y^*(t_i))$ and the fact that $|y(t_i)| \leq L$ (cf. Definition 1.1.4 (ii)). Again,

$$\tilde{v}(X(t_i), \tilde{\bar{x}}(t_i), \tilde{\bar{y}}(t_i)) \leq \tilde{v}(Y(t_i), \bar{x}(t_i), \bar{y}(t_i)) < \tilde{v}(Y(t_i), x(t_i), y(t_i)),$$

and monotonicity follows.

By the induction hypothesis, an optimal strategy for $X(t_{i+1}) = 0$ exists, and therefore the value function $\bar{v}(t_{i+1}, \cdot)$ is bounded from below:

$$\bar{v}(t_{i+1}, \cdot) \geq \bar{v}(t_{i+1}, 0) > -\infty,$$

hence

$$\lim_{\|(x,y)\|\to\infty} \tilde{v}(X(t_i), x, y) = \lim_{|x|\to\infty} \tilde{v}(X(t_i), x, y) = \infty. \tag{1.83}$$

The induction hypothesis implies continuity of $\bar{v}(t_{i+1}, \cdot)$ which in turn implies continuity of $\tilde{v}(\cdot)$. Hence, Equation (1.83) implies the existence of an optimal trading strategy at time t_i. Continuity of $\bar{v}(t_i, \cdot)$ follows directly.

It remains thus to show that optimal liquidation strategies are not market manipulating. We proceed by backward induction again. The case $i = N$ is clear. For $i < N$ and asset position $X(t_i)$, we consider an optimal strategy

$$(x^*(t_i), y^*(t_i)).$$

By the considerations above, we know that

$$\operatorname{sgn}(x^*(t_i)) = \operatorname{sgn}(y^*(t_i)) \quad \text{or} \quad x^*(t_i) \cdot y^*(t_i) = 0.$$

It follows directly from the Bellman Equation (1.81) that

$$\operatorname{sgn}(x^*(t_i)) \neq \operatorname{sgn}(X(t_i)) \quad \text{or} \quad \operatorname{sgn}(y^*(t_i)) \neq \operatorname{sgn}(X(t_i))$$

yield a contradiction as in these cases, orders

$$\tilde{x}(t_i) = \tilde{y}(t_i) = 0$$

result in lower costs. □

Note that neither limited dark pool liquidity nor adverse selection *alone* are sufficient to establish the previous proposition; only the *combination* of the two ensures the desired property.

The assumptions of Proposition 1.5.4 are strong; we leave it for future research to determine tighter necessary and sufficient conditions for the exclusion of profitable market manipulation in markets with dark pools. We only want to remark that our assumptions in Proposition 1.5.2 are not too restrictive for dark pool usage in general: for large initial asset positions $X(t_0)$, the optimal strategy places orders in the dark pool in a non-market manipulating fashion.

2. Portfolio liquidation in continuous time

We transfer the discrete-time trading model of Chapter 1 into a continuous-time trading model in Section 2.1. More precisely, we consider the portfolio liquidation model of Section 1.3 and replace the discrete trades in the primary venue by a continuous trading intensity that can be controlled by the trader. Additionally, we assume that execution of orders in the dark pool is driven by an n-dimensional Poisson process.

The costs of a trading strategy are specified by the continuous-time analog of the Optimization Problem ($\mathrm{OPT}_{\mathrm{dis}}$). Thus we obtain a linear-quadratic stochastic control problem. Due to the liquidation constraint, the value function v of the optimization problem has a singularity at the terminal time. Thus, a verification argument using the HJB equation corresponding to the problem requires non-standard considerations.

We first approximate the liquidation constraint by penalizing non-liquidation by finite end-costs. Heuristic considerations indicate that the value function of the modified optimization problem is quadratic. The corresponding HJB equation suggests that it is given via an initial value problem for a matrix differential equation.

We analyze this initial value problem in Section 2.2. The differential equation is not a Riccati matrix differential equation, and therefore the existing theory for this type of differential equations is not applicable directly. However, via an adequate matrix inequality we can apply a well-known comparison result for Riccati equations and prove existence of the solution of the initial value problem on the whole interval $[0, T]$. Additionally, we are able to obtain upper and lower bounds for this solution in closed form.

These results enable us to solve the modified optimization problem with finite end-costs in Section 2.3. By letting the end-costs tend to infinity, we can transfer the solution to the original optimal liquidation problem in Section 2.4. The key results towards this goal are the bounds obtained in Section 2.2.

In Section 2.5 we determine properties of the optimal liquidation strategy. In Section 2.6 we analyze convergence of the solution of the discrete-time optimization problem of Section 1.3 to the solution of the continuous-time optimization problem obtained in Section 2.4.

2.1. Model description

For a fixed time interval $[0, T]$, we consider the stochastic basis $(\Omega, \mathcal{F}, \mathbb{P}, \mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]})$, where the filtration satisfies the usual conditions.

As in Chapter 1, we investigate a market model where a risk-averse trader with a personal risk-aversion parameter $\alpha \geq 0$ has to liquidate a portfolio $x \in \mathbb{R}^n$ of n assets within a finite trading horizon $[0, T]$. Again, the investor has the possibility to trade simultaneously at a traditional exchange and in a dark pool, facing the trade-off of paying market impact costs in the traditional exchange against uncertain execution in the dark pool.

Here, we replace the discrete orders in the primary venue of Chapter 1 by a trading intensity that can be controlled by the trader at any given point in time $t \in [0, T)$. We also allow for continuous updating of the orders in the dark pool.

In a similar fashion as in Section 1.3, we assume that trading in the traditional exchange generates linear temporary price impact given by a positive definite matrix

$$\Lambda \in \mathbb{R}^{n \times n}, \quad \text{constant in time.}$$

Trading in the dark pool neither generates price impact nor does the trader pay a higher price in the dark pool because of the price impact generated by her trading in the traditional exchange (in contrast to the model in Section 1.5).

The fundamental price at which the assets are traded in the market (provided the large investor does not influence the demand and supply by her actions) is given by an n-dimensional stochastic process

$$\tilde{P} = (\tilde{P}(t))_{t \in [0,T]}.$$

As in Section 1.3, we assume that $\tilde{P}$ is a martingale. By the results of Section 1.3.2, both the optimal strategy and the value function in the case of discrete-time trading are independent of the evolution of the fundamental asset prices. For simplicity and in order to keep the exposition consistent, we will not specify $\tilde{P}$ any further, and the costs of a given trading strategy will not be directly specified via the implementation shortfall (cf. Definition 1.1.5). Rather they will be specified in an analog fashion as in the Optimization Problem ($\mathrm{OPT}_{\mathrm{dis}}$), and the fundamental asset price $\tilde{P}$ will only be relevant through its covariance matrix

$$\Sigma \in \mathbb{R}^{n \times n}, \quad \text{constant in time.}$$

In analogy to the Poisson type execution of orders in the dark pool in the discrete-time model in Sections 1.3 and 1.4, we model trade execution in the dark pool by an n-dimensional Poisson process

$$\pi = (\pi_1, \dots, \pi_n)$$

with intensities

$$\theta_1, \dots, \theta_n \geq 0,$$

respectively. We assume that $\pi_1, \dots, \pi_n$ are independent.

In Section 2.1.1 we specify admissibility of trading strategies and derive a useful moment estimate for the controlled portfolio process that we require for the verification argument in Section 2.3. In Section 2.1.2 we define the cost functional and the optimization problem. Heuristic arguments suggest that the value function of the optimization problem is singular at terminal time T because of the liquidation constraint. Hence, we introduce a modified optimization problem where we drop the liquidation constraint and approximate it by finite end-costs for a portfolio not liquidated by time T as an intermediate step. In Section 2.1.3 we derive a candidate for the value function of the modified optimization problem via a quadratic ansatz and the corresponding HJB Equation. In Section 2.3 we verify that this candidate is indeed the value function of the modified optimization problem. In Section 2.4 we show that if the end-costs tend to infinity, the limit of the candidate value function converges to the value function of the original liquidation problem.

2.1.1. Admissible trading strategies

Let $t \in [0,T)$ be a given point in time and $x \in \mathbb{R}^n$ be the portfolio position of the trader at time t.

The trader has the possibility to trade asset k in the traditional exchange with trading intensity $\xi_k(s)$ at time $s \in [t,T)$ and to place orders $\eta_k(s)$ in the dark pool at time s.

We call a $2n$ - dimensional stochastic process

$$(u(s))_{s\in[t,T)} = (\xi(s),\eta(s))_{s\in[t,T)}$$

a *trading strategy* if ξ is progressively measurable and η is predictable. Given a trading strategy u, the portfolio position at time $s \in [t,T)$ is given by the following controlled stochastic differential equation:

$$\begin{aligned} dX^u(s) &= -\xi(s)ds - \eta(s)d\pi(s) \\ X^u(t) &= x \end{aligned} \tag{2.1}$$

such that the left hand side in (2.1) is well-defined.

For solving the optimization problem we have in mind, we require all trading strategies to fulfill the following conditions.

Definition 2.1.1. *Let $t \in [0,T)$ and $x \in \mathbb{R}^n$ be fixed. Let*

$$u = (u(s))_{s\in[t,T)} = ((\xi(s),\eta(s)))_{s\in[0,T)}$$

be a trading strategy, i.e., ξ is progressively measurable and η is predictable.
(a) We call u an admissible trading strategy *if it fulfills the following conditions.*

(i) The Stochastic Differential Equation (2.1) *possesses a unique solution on* $[t,T)$.

(ii)

$$\begin{aligned} \mathbb{E}_{t,x}\Big[\int_t^T \|\xi(s)\|_2^4 ds\Big] &:= \mathbb{E}\Big[\int_t^T \|\xi(s)\|_2^4 ds \,\Big|\, X^u(t)=x\Big] < \infty, \\ \mathbb{E}_{t,x}\Big[\int_t^T \|\eta(s)\|_2^8 ds\Big] &:= \mathbb{E}\Big[\int_t^T \|\eta(s)\|_2^8 ds \,\Big|\, X^u(t)=x\Big] < \infty. \end{aligned} \tag{2.2}$$

(iii) If $\theta_i = 0$*, then* $\eta_i(s) = 0$ *for all* $s \in [t, T)$*.*

We denote the set of admissible trading strategies by $\tilde{\mathbb{A}}(t, x)$*.*
(b) We call $u \in \tilde{\mathbb{A}}(t, x)$ *an* admissible liquidation strategy *or just* liquidation strategy *if additionally*

(iv)

$$\lim_{s \to T-} X^u(s) = 0 \quad a.s.$$

and denote the set of admissible liquidation strategies by $\mathbb{A}(t, x)$*.*

For convenience, we use from now on the notation in (ii) for expectations conditional of the initial state x of the controlled process at time t.

Remark 2.1.2. *Note that the set of admissible liquidation strategies is not empty. Indeed, let* $t \in [0, T), x \in \mathbb{R}^n$ *and define* $u = (\xi, \eta)$ *by*

$$\begin{aligned} \xi(s) &= \frac{x}{T - t}, \\ \eta(s) &= 0 \end{aligned}$$

for $s \in [t, T)$*. Thus, the dark pool is not used, and the trading intensity in the primary venue is constant. The solution of the Stochastic Differential Equation* (2.1) *is then*

$$X^u(s) = x - \int_t^s \xi(r) dr = \frac{(T - s)x}{T - t}$$

on $[t, T)$*, in particular we have*

$$\lim_{s \to T-} X^u(s) = 0.$$

Remark 2.1.3. *We expect that the stochastic control problems we solve in this chapter and in Chapter 3 are such that the optimal control is of* Markovian form *(see, e.g., the book by Øksendal [2007], Theorem 11.2.3):*

$$u(s) = (\xi(s), \eta(s)) = (\xi(s, X(s)), \eta(s, X(s-))$$

for deterministic functions $\xi, \eta : [t, T) \times \mathbb{R}^n \to \mathbb{R}^n$*.*
The deterministic initial value problem

$$\begin{aligned} X' &= -\xi(\cdot, X) \\ X(t) &= x \end{aligned} \tag{2.3}$$

possesses a unique solution on $[t, T)$ *if*

$$\|\xi(s, y)\| \leq f(s)\|y\| + g(s) \quad on\ [t, T) \times \mathbb{R}^n$$

for $f, g \in C([t,T))$ and $\xi(s,\cdot)$ is locally Lipschitz (e.g., C^1), i.e.,

$$\|\xi(s,y_1) - \xi(s,y_2)\| \leq h(s)\|y_1 - y_2\| \quad \text{for } y_1, y_2 \in \mathbb{R}^n, s \in [t,T),$$

where h is locally bounded (e.g., by Peano's existence theorem and Gronwall's inequality).

Let $\xi : [t,T) \times \mathbb{R}^n \to \mathbb{R}^n$ fulfill these conditions and let $\eta : [t,T) \times \mathbb{R}^n \to \mathbb{R}^n$. We can pathwise construct the solution of the Stochastic Differential Equation (2.1) inductively by interlacing the jumps (see, e.g., Applebaum [2004], Example 1.3.13): as the n Poisson processes are independent, they jump at distinct times almost surely. Let $(\tau_i)_{i\geq 1}$ be the jump times of π such that

$$t =: \tau_0 < \tau_1 < \ldots \text{ almost surely,}$$

and let X be the solution of the Initial Value Problem (2.3) on $[\tau_i, \tau_{i+1} \wedge T)$ with initial value $x = X(\tau_i)$ for $i \in \mathbb{N}$ such that $\tau_i \leq T$. For $\tau_{i+1} \leq T$ and $\Delta\pi_k(\tau_{i+1}) > 0$, we set

$$X(\tau_{i+1}) := X(\tau_{i+1}-) - \eta_k(\tau_{i+1}, X(\tau_{i+1}-))e_k,$$

where e_k is the k^{th} unit vector.

Definition 2.1.1 implies the following moment estimate for the controlled process. We need this result later for the proof of the verification theorem in Section 2.3.2.

Proposition 2.1.4. *Let $t \in [0,T)$, $x \in \mathbb{R}^n$ and $u \in \tilde{\mathbb{A}}(t,x)$. Then*

$$\mathbb{E}_{t,x}\Big[\sup_{t\leq s\leq T} \|X^u(s)\|_2^4\Big] < \infty,$$

in particular

$$\mathbb{E}_{t,x}\Big[\int_t^T \|X^u(s)\|_2^4 ds\Big] < \infty.$$

Proof. Let $s \in [t,T]$. Then

$$\begin{aligned}
\|X^u(s)\|_2^4 &\leq \Big(\|x\|_2 + \Big\|\int_t^s \xi(r)dr\Big\|_2 + \Big\|\int_t^s \eta(r)d\pi(r)\Big\|_2\Big)^4 \\
&\leq 3^3\Big(\|x\|_2^4 + \Big\|\int_t^s \xi(r)dr\Big\|_2^4 + \Big\|\int_t^s \eta(r)d\pi(r)\Big\|_2^4\Big) \\
&\leq 27\Big(\|x\|_2^4 + (s-t)^3\int_t^s \|\xi(r)\|_2^4 dr + \Big\|\int_t^s \eta(r)d\pi(r)\Big\|_2^4\Big), \qquad (2.4)
\end{aligned}$$

where Inequality (2.4) follows from a multidimensional version of Jensen's inequality (see, e.g., Kallenberg [2002], Lemma 3.5). Note first that by Definition 2.1.1 (ii),

$$\mathbb{E}_{t,x}\Big[(s-t)^3 \int_t^s \|\xi(r)\|_2^4 dr\Big] \le (T-t)^3 \mathbb{E}_{t,x}\Big[\int_t^T \|\xi(r)\|_2^4 dr\Big] < K \tag{2.5}$$

for a constant K independent of s. Let us now consider the compensated Poisson processes

$$M_i(s) := \pi_i(s) - \theta_i s, \quad i = 1, \dots, n.$$

Then by Hölder's inequality,

$$\begin{aligned}
\Big|\int_t^s \eta_i(r) d\pi_i(r)\Big|^4 &= \Big|\sum_{r,\ \Delta\pi_i(r)\neq 0} \eta_i(r)\Big|^4 \\
&\le \big(\pi_i(s) - \pi_i(t)\big)^3 \int_t^s |\eta_i(r)|^4 d\pi_i(r) \\
&= \big(\pi_i(s) - \pi_i(t)\big)^3 \Big(\int_t^s |\eta_i(r)|^4 dM_i(r) + \theta_i \int_t^s |\eta_i(r)|^4 dr\Big).
\end{aligned}$$

We obtain

$$\begin{aligned}
&\mathbb{E}_{t,x}\Big[\big(\pi_i(s) - \pi_i(t)\big)^3 \int_t^s |\eta_i(r)|^4 dM_i(r)\Big] \\
&\quad\le \mathbb{E}_{t,x}\Big[\big(\pi_i(s) - \pi_i(t)\big)^6\Big]^{\frac{1}{2}} \mathbb{E}_{t,x}\Big[\Big(\int_t^s |\eta_i(r)|^4 dM_i(r)\Big)^2\Big]^{\frac{1}{2}} && (2.6) \\
&\quad\le \mathbb{E}_{t,x}\Big[\pi_i(T-t)^6\Big]^{\frac{1}{2}} \mathbb{E}_{t,x}\Big[\theta_i \int_t^T |\eta_i(r)|^8 dr\Big]^{\frac{1}{2}} && (2.7) \\
&\quad< K_i && (2.8)
\end{aligned}$$

for a constant K_i independent of s. Inequality (2.6) follows from Hölder's inequality, Inequality (2.7) follows from Itô's isometry (note that $\langle M_i\rangle(s) = \theta_i s$) and Inequality (2.8) follows from Definition 2.1.1 (ii) and the fact that Poisson distributed random variables have finite moments. Therefore, by Definition 2.1.1 (ii) and Hölder's inequality (recall that e_i denotes the i^{th} unit vector),

$$\begin{aligned}
\mathbb{E}_{t,x}\Big[\Big\|\int_t^s \eta(r) d\pi(r)\Big\|_2^4\Big] &= \mathbb{E}_{t,x}\Big[\Big\|\sum_{i=1}^n e_i \int_t^s \eta_i(r) d\pi_i(r)\Big\|_2^4\Big] \\
&\le n^3 \sum_{i=1}^n \mathbb{E}_{t,x}\Big[\Big|\int_t^s \eta_i(r) d\pi_i(r)\Big|^4\Big]
\end{aligned}$$

$$< \tilde{K} \tag{2.9}$$

for a constant $\tilde{K}$ independent of s. Combining Inequalities (2.4), (2.5) and (2.9) completes the proof. □

2.1.2. Cost functional

In the discrete-time setting of Section 1.3.2 we showed that both the optimal strategy and the value function are independent of the evolution of the fundamental asset price. We define the costs of continuous-time admissible liquidation strategies $u \in \mathbb{A}(t,x)$ analogously to the costs of discrete-time strategies in the Optimization Problem (OPT_{dis}).

Let $t \in [0,T)$, $x \in \mathbb{R}^n$. Then a liquidation strategy $u \in \mathbb{A}(t,x)$ yields the trading costs

$$J(t,x,u) := \mathbb{E}_{t,x}\Big[\int_t^T f(\xi(s), X^u(s))ds\Big],$$

where $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is given by

$$f(\xi,x) := \xi^\top \Lambda \xi + \alpha x^\top \Sigma x. \tag{2.10}$$

The first term of the right hand side of Equation (2.10) refers to the linear price impact costs generated by trading in the traditional market. The second term refers to quadratic risk costs penalizing slow liquidation. The trader aims to minimize her trading costs and considers the following optimization problem:

$$v(t,x) := \inf_{u \in \mathbb{A}(t,x)} J(t,x,u). \tag{OPT}$$

Note that the optimization problem is well-defined and the value function satisfies

$$v(t,x) < \infty$$

as for constant liquidation exclusively in the primary exchange, we have

$$v(t,x) \le J(t,x,u) = \int_t^T \Big(\frac{1}{(T-t)^2} x^\top \Lambda x + \frac{(T-s)^2}{(T-t)^2}\alpha x^\top \Sigma x\Big) ds < \infty.$$

Because of the liquidation constraint (cf. Definition 2.1.1 (iv)), we expect the value function to fulfill

$$\lim_{s \to T-} v(s,x) = \begin{cases} 0 & \text{if } x = 0 \\ \infty & \text{else,} \end{cases}$$

i.e., v has a singularity at the terminal time T. Because of this singularity, non-standard considerations are necessary for solving the Optimization Problem (OPT) via a verification argument using the HJB equation.

As an intermediate step, we hence weaken the liquidation constraint by allowing for all strategies $u \in \tilde{\mathbb{A}}(t,x)$ and penalizing non-liquidation by finite end-costs. More precisely, for $l > 0$ and $u = (\xi, \eta) \in \tilde{\mathbb{A}}(t,x)$, we define the following cost functional

$$\tilde{J}(l,t,x,u) := \mathbb{E}_{t,x}\Big[\int_t^T f(\xi(s), X^u(s))ds + l \cdot X^u(T)^\top X^u(T)\Big].$$

The resulting optimization problem is

$$\tilde{v}(l,t,x) := \inf_{u \in \tilde{\mathbb{A}}(t,x)} \tilde{J}(l,t,x,u). \tag{$\widetilde{\text{OPT}}$}$$

In the following, we solve the modified Optimization Problem ($\widetilde{\text{OPT}}$) first. Later we show that the solution of the Optimization Problem ($\widetilde{\text{OPT}}$) converges to the solution of the original Optimization Problem (OPT) as $l \to \infty$ (Section 2.4).

2.1.3. Hamilton-Jacobi-Bellman equation

In this section we derive a candidate for the value function of the Optimization Problem ($\widetilde{\text{OPT}}$). Heuristic considerations suggest that it should satisfy the following HJB equation (see, e.g., the book by Øksendal and Sulem [2007]):

$$\begin{aligned}\frac{\partial w}{\partial t}(t,x) &= \sup_{u=(\xi,\eta)\in\mathbb{R}^n\times\mathbb{R}^n}\Big[\sum_{i=1}^n \theta_i\big(w(t,x) - w(t, x-\eta^\top e_i)\big) + \nabla_x w(t,x)\xi - f(\xi,x)\Big] \\ w(T,x) &= l x^\top x.\end{aligned} \tag{$\widetilde{\text{HJB}}$}$$

There is reason to believe that the value function is quadratic in x (cf. Section 1.3). Given that the above guesses are correct, the following lemma provides candidates both for the value function and for the optimal strategy.

Lemma 2.1.5. *Let $w : \mathbb{R}_+ \times \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}$ be defined by*

$$w(l,t,x) := x^\top C(l,t)x$$

for positive definite matrices

$$C(l,t) = \big(c_{i,j}(l,t)\big)_{i,j=1,\dots,n} \in \mathbb{R}^{n\times n}, \quad t \in [0,T],$$

such that $c_{i,j}(l,\cdot) \in C^1([0,T])$ for $i,j = 1,\dots,n$.

Let $h : \mathbb{R}_+ \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be given by

$$h(l,t,x,\xi,\eta) := \sum_{i=1}^n \theta_i\big(w(t,x) - w(t,x-\eta^\top e_i)\big) + \nabla_x w(l,t,x)\xi - f(\xi,x)$$

$$= \sum_{i=1}^{n} \theta_i \Big(x^\top C(l,t)x - (x-\eta_i e_i)^\top C(l,t)(x-\eta_i e_i) \Big) \\ + 2x^\top C(l,t)\xi - \xi^\top \Lambda \xi - \alpha x^\top \Sigma x. \tag{2.11}$$

For fixed $l > 0$, $t \in [0,T]$ and $x \in \mathbb{R}^n$, $h(l,t,x,\cdot,\cdot)$ attains its maximum at

$$\xi^* := \xi^*(l,t,x) := \Lambda^{-1} C(l,t)x, \tag{2.12}$$

$$\eta^* := \eta^*(l,t,x) := \bar{C}(l,t)C(l,t)x, \tag{2.13}$$

where

$$\bar{C}(l,t) := \operatorname{diag}\Big(\frac{1}{c_{i,i}(l,t)}\Big). \tag{2.14}$$

Moreover,

$$h(l,t,x,\xi^*,\eta^*) = x^\top C(l,t)\Lambda^{-1}C(l,t)x + x^\top C(l,t)\tilde{C}(l,t)C(l,t)x - \alpha x^\top \Sigma x, \tag{2.15}$$

where

$$\tilde{C}(l,t) := \operatorname{diag}\Big(\frac{\theta_i}{c_{i,i}(l,t)}\Big).$$

If $\theta_i > 0$ for all $i = 1,\dots,n$, the maximum is obtained uniquely at (ξ^*,η^*) as in Equations (2.12) *and* (2.13). *If there exist $i_1,\dots,i_k \in \{1,\dots,n\}$ such that $\theta_{i_j} = 0$ $(j = 1,\dots,k)$, then η_{i_j} can be chosen arbitrarily without changing the value of h. Up to arbitrary choices of η_{i_j}, the maximizer given in Equations* (2.12) *and* (2.13) *is unique.*

Proof. We have

$$h(l,t,x,\xi,\eta) = -\sum_{i=1}^{n} \theta_i c_{i,i}(l,t)\Big(\eta_i - e_i^\top \frac{1}{c_{i,i}(l,t)} C(l,t)x\Big)^2 \\ - (\xi - \Lambda^{-1}C(l,t)x)^\top \Lambda (\xi - \Lambda^{-1}C(l,t)x) \\ + x^\top C(l,t)\tilde{C}(l,t)C(l,t)x + x^\top C(l,t)\Lambda^{-1}C(l,t)x - \alpha x^\top \Sigma x. \tag{2.16}$$

If $\theta_i > 0$ for all $i = 1,\dots,n$, the term in Equation (2.16) is maximal if and only if $\xi = \xi^*$ and $\eta = \eta^*$ for ξ^* and η^* as in Equations (2.12) and (2.13), respectively. If there exist $i_1,\dots,i_k \in \{1,\dots,n\}$ such that $\theta_{i_j} = 0$ $(j = 1,\dots,k)$, η_{i_j} can be chosen arbitrarily without changing the value of h.

Plugging (ξ^*,η^*) into Equation (2.11), we obtain Equation (2.15). □

We can directly deduce a candidate for the value function via the solution of the Matrix Differential Equation (2.17).

Corollary 2.1.6. *Let $l > 0$ and assume that the initial value problem for a matrix differential equation*

$$\begin{aligned} C' &= C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha \Sigma \\ C(T) &= lI \end{aligned} \tag{2.17}$$

possesses a positive definite solution $C(l,t)$ on $[0,T]$. Then

$$w(l,t,x) := x^\top C(l,t)x$$

satisfies the HJB Equation $(\widetilde{\text{HJB}})$ *with maximizer $u^* = (\xi^*, \eta^*)$ for*

$$\begin{aligned}\xi^* &:= \xi^*(l,t,x) := \Lambda^{-1}C(l,t)x,\\ \eta^* &:= \eta^*(l,t,x) := \bar{C}(l,t)C(l,t)x\end{aligned}$$

(as in Equations (2.12) *and* (2.13)*).*

2.2. Riccati matrix differential equations and inequalities

Corollary 2.1.6 suggests that we have to study the following initial value problem for a matrix differential equation

$$\begin{aligned}C' &= C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha\Sigma\\ C(T) &= lI,\end{aligned} \tag{2.18}$$

where

$$\tilde{C} := \operatorname{diag}\Big(\frac{\theta_i}{c_{i,i}}\Big). \tag{2.19}$$

It is not immediately clear that the Initial Value Problem (2.18) possesses a positive definite solution on the whole interval $[0,T]$ for $n \geq 2$. For $n = 1$, it reduces to

$$\begin{aligned}C' &= \frac{C^2}{\Lambda} + \theta_1 C - \alpha\Sigma\\ C(T) &= l.\end{aligned} \tag{2.20}$$

This is an initial value problem for a scalar Riccati differential equation with constant coefficients, whose unique solution is explicitly known (cf. Section 2.2.2) and exists on the whole interval $[0,T]$. For $n \geq 2$, the second summand in the matrix differential equation

$$C^\top \tilde{C} C$$

is in general not linear (or quadratic), and (2.18) is not a Riccati matrix differential equation. Furthermore, a closed form solution for the corresponding initial value problem is not known, and the existing theory about Riccati matrix differential equations is not applicable directly.

It turns out that appropriate upper and lower bounds for the non-linear term $C^\top \tilde{C} C$ transform to lower and upper bounds for the solution of the Matrix Initial Value Problem (2.18) and yield existence and positive definiteness of the solution on the whole interval $(-\infty, T]$ (but not on $\mathbb{R}$ in general). To this end, we require a version of a well-known comparison result for matrix Riccati differential equations, which we state and prove in Section 2.2.1. The main step is thus to obtain adequate matrix inequalities which enable us to transfer these results to the Initial Value Problem (2.18). Before we execute this task in Section 2.2.3, we treat scalar Riccati differential equations with constant coefficients, which can be solved explicitly in Section 2.2.2. Finally, we combine the results of Sections 2.2.1, 2.2.2 and 2.2.3 in order to prove the existence of the solution of (2.18) on $(-\infty, T]$ in Section 2.2.4. We are also able to compute relatively simple upper and lower bounds for the solution in closed form. These bounds will be a main building block for the verification arguments in Section 2.3 and Section 2.4.

Before we proceed, we introduce the following notations.

Notation 2.2.1. *(i) For symmetric matrices A and B we say*

$$A > 0 \quad (A \geq 0)$$

if A is positive (nonnegative) definite. We say

$$A > B \quad (A \geq B)$$

if $A - B$ is positive (nonnegative) definite.

(ii) We denote the smallest and the largest eigenvalues of a real-symmetric matrix A by $a_{\min}$ and $a_{\max}$, respectively.

(iii) We define the following nonnegative definite matrix:

$$D := \sqrt{\Lambda^{-1}}\Sigma\sqrt{\Lambda^{-1}}.$$

2.2.1. Riccati matrix differential equations

In this section we prove a well-known comparison result for the solutions of Riccati matrix differential equations. We only state and prove the result in the generality necessary for the application we have in mind. For a general approach to Riccati matrix differential equations, see, e.g., the book by Reid [1972].

Theorem 2.2.2. *Let $A(t), B_P(t), C_P(t), B_Q(t), C_Q(t) \in \mathbb{R}^{n\times n}$ be piecewise continuous on $\mathbb{R}$. Furthermore, let $B_P(t), C_P(t), B_Q(t), C_Q(t)$ $(t \in \mathbb{R})$ and $S_P, S_Q \in \mathbb{R}^{n\times n}$ be symmetric.*

(i) Let $t_0 < t_1 \leq \infty$ and

$$S_P \leq S_Q, \quad 0 \leq B_Q(\cdot) \leq B_P(\cdot), \quad C_P(\cdot) \leq C_Q(\cdot) \tag{2.21}$$

on $[t_0, t_1)$. Assume that the initial value problem

$$\begin{aligned} P' &= -A^\top P - PA - PB_P P + C_P \\ P(t_0) &= S_P \end{aligned} \tag{2.22}$$

possesses a solution P on $[t_0, t_1)$. Then the initial value problem

$$\begin{aligned} Q' &= -A^\top Q - QA - QB_Q Q + C_Q \\ Q(t_0) &= S_Q \end{aligned} \tag{2.23}$$

possesses a solution Q on $[t_0, t_1)$ and

$$P(t) \leq Q(t) \quad \textit{on } [t_0, t_1).$$

(ii) Let $t_0 > t_2 \geq -\infty$ and

$$S_Q \leq S_P, \quad 0 \leq B_Q(\cdot) \leq B_P(\cdot), \quad C_P(\cdot) \leq C_Q(\cdot)$$

on $(t_2, t_0]$. Assume that the Initial Value Problem (2.22) *possesses a solution P on $(t_2, t_0]$. Then the Initial Value Problem* (2.23) *possesses a solution Q on $(t_2, t_0]$ and*

$$P(t) \geq Q(t) \quad \textit{on } (t_2, t_0].$$

Proof. Note first that uniqueness, local existence and symmetry of the solutions $P(\cdot)$ and $Q(\cdot)$ follow from the Picard-Lindelöf theorem. We only prove the first assertion. Straightforward modifications yield a proof of the second assertion.

Let

$$\tau := \sup\{t \in [t_0, t_1) | Q \text{ exists on } [t_0, t)\}$$

and Ψ be the fundamental matrix of

$$x' = (A + B_Q Q)x \quad \text{on } [t_0, \tau) \quad \text{with } \Psi(t_0) = I.$$

In the first step of the proof we show that

$$\Psi(t)^\top (Q(t) - P(t))\Psi(t) \geq 0 \quad \text{on } [t_0, \tau),$$

which implies by regularity of Ψ on $[t_0, \tau)$ that

$$P(t) \leq Q(t) \quad \text{on } [t_0, \tau). \tag{2.24}$$

To this end, we compute

$$\begin{aligned} &\frac{d}{dt}\Psi(t)^\top (Q(t) - P(t))\Psi(t) \\ &\quad = \Psi(t)^\top (A(t) + B_Q(t)Q(t))^\top (Q(t) - P(t))\Psi(t) \\ &\qquad + \Psi(t)^\top (Q(t) - P(t))(A(t) + B_Q(t)Q(t))\Psi(t) + \Psi(t)^\top (Q'(t) - P'(t))\Psi(t) \\ &\quad = \Psi(t)^\top (C_Q(t) - C_P(t) + Q(t)B_Q(t)Q(t) + P(t)B_P(t)P(t) \\ &\qquad - Q(t)B_Q(t)P(t) - P(t)B_Q(t)Q(t))\Psi(t) \end{aligned}$$

$$
\begin{aligned}
&= \Psi(t)^\top \big(C_Q(t) - C_P(t) + (Q(t) - P(t))B_Q(t)(Q(t) - P(t)) \\
&\qquad + P(t)(B_P - B_Q)P(t)\big)\Psi(t) \\
&\geq 0
\end{aligned}
$$

by (2.21). We deduce

$$
\begin{aligned}
&\Psi(t)^\top (Q(t) - P(t))\Psi(t) \\
&\qquad = \Psi(t_0)^\top (Q(t_0) - P(t_0))\Psi(t_0) + \int_{t_0}^{t} \frac{d}{ds}\Psi(s)^\top (Q(s) - P(s))\Psi(s)ds \geq 0
\end{aligned}
$$

on $[t_0, \tau)$ by (2.21) and the inequality above as desired.

Let now Φ be the fundamental matrix of

$$
x' = Ax \quad \text{on } \mathbb{R} \quad \text{with } \Phi(t_0) = I
$$

and define

$$
Z(t) := (\Phi^{-1}(t))^\top \Big(S_Q + \int_{t_0}^{t} \Phi(s)^\top C_Q(s)\Phi(s)ds\Big)\Phi^{-1}(t).
$$

In the next step of the proof, we show that

$$
\Phi(t)^\top (Z(t) - Q(t))\Phi(t) \geq 0 \text{ on } [t_0, \tau),
$$

hence by regularity of Φ on $[t_0, \tau)$,

$$
Q(t) \leq Z(t) \text{ on } [t_0, \tau). \tag{2.25}
$$

We note first that

$$
\Phi(t)^\top Z(t)\Phi(t) = S_Q + \int_{t_0}^{t} \Phi(s)^\top C_Q(s)\Phi(s)ds
$$

and compute

$$
\begin{aligned}
&\frac{d}{dt}\Phi(t)^\top (Z(t) - Q(t))\Phi(t) \\
&\quad = \Phi(t)^\top C_Q(t)\Phi(t) - \Phi'(t)^\top Q(t)\Phi(t) - \Phi(t)^\top Q'(t)\Phi(t) - \Phi(t)^\top Q(t)\Phi'(t) \\
&\quad = \Phi(t)^\top Q(t)B_Q(t)Q(t)\Phi(t) \\
&\quad \geq 0
\end{aligned}
$$

by (2.21). Similarly as before (note that $Z(t_0) - Q(t_0) = S_Q - S_Q = 0$),

$$
\Phi(t)^\top (Z(t) - Q(t))\Phi(t) = \int_{t_0}^{t} \Phi(s)^\top Q(s)B_Q(s)Q(s)\Phi(s)ds \geq 0
$$

on $[t_0, \tau)$, as required.

To finish the proof, we need to show that $\tau = t_1$. Let us assume that $\tau < t_1$. Since by our assumption on P, P and Z are continuous on the closed interval $[t_0, \tau]$, we have by Inequalities (2.24) and (2.25) that

$$-cI \leq P(t) \leq Q(t) \leq Z(t) \leq cI \quad \text{on } [t_0, \tau) \tag{2.26}$$

for some constant $c > 0$. Thus for $t \in [t_0, \tau)$,

$$\begin{aligned} Q(t) = Q(t_0) + &\int_{t_0}^{t} \big(-A^\top Q - QA - QB_Q Q + C_Q \big)(s) ds \\ \longrightarrow S_Q + &\int_{t_0}^{\tau} \big(-A^\top Q - QA - QB_Q Q + C_Q \big)(s) ds \quad \text{as } t \to \tau- \end{aligned}$$

since the integrand is bounded. By the existence theorem of Picard-Lindelöf, we obtain existence of Q on $[t_0, \tau + \epsilon)$ for some $\epsilon > 0$, contradicting the definition of τ. □

2.2.2. Scalar Riccati differential equations

We will later use the following scalar Riccati differential equation with constant coefficients, which can be solved explicitly (cf. the Initial Value Problem (2.20)). Let us consider the problem

$$\begin{aligned} y' &= y^2 + ay - b \\ y(T) &= c, \end{aligned} \tag{2.27}$$

where $a, b \geq 0$, $c > 0$, $b < c^2 + ac$. If we substitute $z(t) := y(t) + \frac{a}{2}$, we obtain the initial value problem

$$\begin{aligned} z' &= z^2 - d \\ z(T) &= c + \frac{a}{2}, \end{aligned} \tag{2.28}$$

where $d := \frac{a^2}{4} + b \geq 0$. Let first $d > 0$. Noting that $\coth' = 1 - \coth^2$, we obtain that the Initial Value Problem (2.28) is solved by

$$z(t) = \sqrt{d} \coth\big(\sqrt{d}(T - t) + \kappa\big)$$

on $(-\infty, T]$, where (note that $c + \frac{a}{2} > \sqrt{d}$)

$$\kappa = \operatorname{arcoth}\Big(\frac{c + \frac{a}{2}}{\sqrt{d}}\Big) > 0.$$

Thus, the Initial Value Problem (2.27) is solved by

$$y(t) = \sqrt{d} \coth\left(\sqrt{d}(T-t) + \kappa\right) - \frac{a}{2} \tag{2.29}$$

on $(-\infty, T]$. Note also that

$$\lim_{t\to-\infty} y(t) = \sqrt{d} - \frac{a}{2} > c > 0.$$

For $d = 0$ (i.e., $a = b = 0$), we obtain $z' = z^2$, $z = y$ and the Initial Value Problem (2.27) is solved by

$$y(t) = \frac{1}{T - t + \frac{1}{c}}. \tag{2.30}$$

The following simple upper and lower bounds for the solution of the Initial Value Problem (2.27) will later prove to be convenient.

Corollary 2.2.3. *Let $d > 0$ and y be the solution of the Initial Value Problem* (2.27), *i.e., y is as in Equation* (2.29).

(i) Let $t \in (-\infty, T]$, then

$$y(t) \geq \frac{1}{T - t + \frac{1}{c+a/2}} - \frac{a}{2}. \tag{2.31}$$

(ii) Let $t \in (-\infty, T]$ and $c > \sqrt{d}$, then

$$y(t) \leq \frac{1}{T - t + \frac{1}{c-\sqrt{d}+a/2}} + \sqrt{d} - \frac{a}{2}. \tag{2.32}$$

Proof. (i) As $d > 0$, we have that the solution z of the Initial Value Problem (2.28) fulfills

$$z(t) \geq \frac{1}{T - t + \frac{1}{c+a/2}} \text{on } (-\infty, T] \tag{2.33}$$

(cf. Theorem 2.2.2 (ii)). Equation (2.31) follows directly.

(ii) Let

$$h(t) := \frac{1}{T - t + \frac{1}{c-\sqrt{d}+a/2}} + \sqrt{d}.$$

Then $h(T) = c + \frac{a}{2}$ and $h(t) > \sqrt{d}$ on $(-\infty, T]$ (recall $c + \frac{a}{2} > \sqrt{d}$). Thus,

$$h'(t) = (h(t) - \sqrt{d})^2 = h(t)^2 - 2\sqrt{d}h(t) + d < h(t)^2 - d. \tag{2.34}$$

By Theorem 2.2.2 (ii), $h(t) \geq z(t)$, and Equation (2.32) follows.

□

2.2.3. Matrix inequalities

The matrix differential equation in (2.18) is not a Riccati type matrix differential equation and therefore the existing theory is not applicable; however, some of the methods are. The main aim of this section is to bound the solution of (2.18) from below and from above by systems that are Riccati differential equations and therefore easier to handle.

The inconvenient part of (2.18) is the non-linear term

$$C\tilde{C}C. \tag{2.35}$$

For $C > 0$, we have

$$0 \leq C\tilde{C}C.$$

Applying the results from the previous sections, we can use this lower bound to construct an upper bound for the solution of the Initial Value Problem (2.18) in Section 2.2.4. This upper bound is the solution of a matrix Riccati differential equation which is known in closed form. On the other hand, it turns out that we need an upper bound of (2.35) for obtaining a lower bound for the solution of the Initial Value Problem (2.18). Such a bound is obtained in Corollary 2.2.5. It is a consequence of the matrix inequality stated in the following theorem. This theorem is hence a vital component for the solution of the Optimization Problems (OPT) and $(\widetilde{\text{OPT}})$.

Theorem 2.2.4. *Let $C = (c_{i,j})_{i,j=1\dots,n} \in \mathbb{R}^{n\times n}$ be a positive definite matrix and $\theta_i > 0$, $i = 1, \dots, n$. Then*

$$C \leq \theta \operatorname{diag}\left(\frac{c_{i,i}}{\theta_i}\right) = \theta \tilde{C}^{-1}, \tag{2.36}$$

where

$$\theta := \theta(n) := \sum_{i=1}^{n} \theta_i.$$

Proof. We prove the inequality by induction on n. It is clear for $n = 1$ with equality in (2.36).

Let now $n \geq 1$ and $C = (c_{i,j})_{i,j=1,\dots,n+1} \in \mathbb{R}^{(n+1)\times(n+1)}$ be positive definite. Define $C_n \in \mathbb{R}^{n\times n}$ and $c \in \mathbb{R}^n$ such that

$$C = \begin{pmatrix} C_n & c \\ c^\top & c_{n+1,n+1} \end{pmatrix}.$$

For $z = (x, y)^\top \in \mathbb{R}^n \times \mathbb{R}$, $z \neq 0$, we have

$$\begin{aligned} z^\top C z &= x^\top C_n x + 2x^\top c y + c_{n+1,n+1} y^2, \\ z^\top \operatorname{diag}\left(\frac{c_{i,i}}{\theta_i}\right) z &= x^\top \operatorname{diag}\left(\frac{c_{i,i}}{\theta_i}\right) x + \frac{c_{n+1,n+1}}{\theta_{n+1}} y^2. \end{aligned}$$

Note that we use $\operatorname{diag}\left(\frac{c_{i,i}}{\theta_i}\right)$ both for the diagonal $n \times n$ - matrix with $\frac{c_{1,1}}{\theta_1}, \dots, \frac{c_{n,n}}{\theta_n}$ in the diagonal and for the respective diagonal $(n + 1) \times (n + 1)$ - matrix. Which one we refer to is always clear from the context.

$$
\begin{aligned}
&z^\top\Big(\theta(n+1)\operatorname{diag}\Big(\frac{c_{i,i}}{\theta_i}\Big)-C\Big)z\\
&\quad=\theta(n+1)x^\top\operatorname{diag}\Big(\frac{c_{i,i}}{\theta_i}\Big)x+\frac{\theta(n+1)}{\theta_{n+1}}c_{n+1,n+1}y^2\\
&\qquad-x^\top C_n x-2x^\top cy-c_{n+1,n+1}y^2\\
&\quad=\underbrace{x^\top\Big(\theta(n)\operatorname{diag}\Big(\frac{c_{i,i}}{\theta_i}\Big)-C_n\Big)x^\top}_{\geq 0\text{ by the induction hypothesis}}\\
&\qquad\underbrace{+\,\theta_{n+1}x^\top\operatorname{diag}\Big(\frac{c_{i,i}}{\theta_i}\Big)x-2x^\top cy+\frac{\theta(n)}{\theta_{n+1}}c_{n+1,n+1}y^2}_{(*)}\,.
\end{aligned}
$$

It is thus sufficient to show that $(*)\geq 0$.

$$
\begin{aligned}
(*)=&\Big(x-\frac{1}{\theta_{n+1}}\operatorname{diag}\Big(\frac{\theta_i}{c_{i,i}}\Big)cy\Big)^\top\theta_{n+1}\operatorname{diag}\Big(\frac{c_{i,i}}{\theta_i}\Big)\Big(x-\frac{1}{\theta_{n+1}}\operatorname{diag}\Big(\frac{\theta_i}{c_{i,i}}\Big)cy\Big)\\
&+\Big(c_{n+1,n+1}\frac{\theta(n)}{\theta_{n+1}}-c^\top\frac{\operatorname{diag}\big(\frac{\theta_i}{c_{i,i}}\big)}{\theta_{n+1}}c\Big)y^2, \qquad (2.37)
\end{aligned}
$$

where the first term in Equation (2.37) is nonnegative as C_n (and therefore $\operatorname{diag}\big(\frac{c_{i,i}}{\theta_i}\big)$) is positive definite and $\theta_{n+1}>0$. We will now use the following property for positive definite matrices A and B (see, e.g., the book by Horn and Johnson [1985], Corollary 7.7.4):

$$\text{If } 0<A<B,\quad\text{then } 0<B^{-1}<A^{-1}. \qquad (2.38)$$

Thus,

$$0<\frac{1}{\theta(n)}\operatorname{diag}\Big(\frac{\theta_i}{c_{i,i}}\Big)\leq C_n^{-1} \qquad (2.39)$$

by the induction hypothesis. Moreover,

$$C\begin{pmatrix}I_{n\times n} & -C_n^{-1}c\\ 0 & 1\end{pmatrix}=\begin{pmatrix}C_n & 0\\ c^\top & c_{n+1,n+1}-c^\top C_n^{-1}c\end{pmatrix}$$

and hence

$$c_{n+1,n+1}-c^\top C_n^{-1}c=\frac{\det C}{\det C_n}>0. \qquad (2.40)$$

Finally,

$$c_{n+1,n+1}-\frac{1}{\theta(n)}c^\top\operatorname{diag}\Big(\frac{\theta_i}{c_{i,i}}\Big)c\overset{(2.39)}{\geq}c_{n+1,n+1}-c^\top C_n^{-1}c\overset{(2.40)}{>}0,$$

and the second term in (2.37) is positive. $\square$

Note that the same line of reasoning establishes strict inequality in Inequality (2.36) for $n \geq 2$. We can now derive the desired upper bound for (2.35).

Corollary 2.2.5. *For a positive definite matrix $C \in \mathbb{R}^{n\times n}$ and $\theta_i \geq 0$, $i = 1, \ldots, n$, we have*

$$C\tilde{C}C \leq \theta C. \tag{2.41}$$

Proof. Let first $\theta_i > 0$ for all $i = 1, \ldots, n$. By Theorem 2.2.4, we have

$$0 < C \leq \theta \tilde{C}^{-1}$$

and therefore (cf. Property (2.38))

$$\tilde{C} \leq \theta C^{-1}.$$

As C is regular, this is equivalent to Inequality (2.41). If $\theta_{i_1} = \cdots = \theta_{i_k} = 0$ for some $i_1, \ldots, i_k \in \{1, \ldots, n\}$, we set $\tilde{\theta}_{i_l} = \theta_{i_l} + \epsilon$ and let ϵ tend to 0 in Inequality (2.41). $\square$

We close the section with two surprisingly simple matrix inequalities that are special cases of Theorem 2.2.4. The first one proves useful for applications later.

Corollary 2.2.6. *For a positive definite matrix $C \in \mathbb{R}^{n\times n}$, we have*

$$C \leq n \operatorname{diag}\left(c_{i,i}\right)$$

and

$$C \leq \operatorname{tr}(C) I.$$

Proof. The inequalities follow from Theorem 2.2.4 for $\theta_i = 1$ respectively $\theta_i = c_{i,i}$ ($i = 1, \ldots, n$). $\square$

2.2.4. The key matrix differential equation

The matrix inequalities of Section 2.2.3 enable us to apply Theorem 2.2.2 to the Matrix Initial Value Problem (2.18) such that we can prove existence of a solution C of (2.18) on the whole interval $(-\infty, T]$ and at the same time construct upper and lower bounds for C. For technical reasons, we will rather bound

$$\sqrt{\Lambda^{-1}} C \sqrt{\Lambda^{-1}}$$

though. The bounds are constructed in terms of multiples of the identity matrix and hence commute with all matrices. Therefore, they transfer to bounds of C directly by multiplying them with Λ.

In the following Lemma we solve two Riccati matrix initial value problems in closed form. These solutions are proven to be the desired bounds in Theorem 2.2.8 below. This is the key step for solving the Optimization Problems (OPT) and ($\widetilde{\text{OPT}}$). First, we obtain a well-defined candidate for the value function of ($\widetilde{\text{OPT}}$). Secondly, the bounds of C are needed for many important arguments of the corresponding verification proof in Section 2.3. Finally, the bounds enable us to transfer the solution of ($\widetilde{\text{OPT}}$) to a well-defined candidate for the solution of the Optimization Problem (OPT) by letting l tend to infinity in Section 2.4. The limits of the upper and lower bounds are then important for the verification proof.

Lemma 2.2.7. *Let $\theta_i \geq 0$ for $i = 1, \dots, n$, $\theta = \sum_{i=1}^{n} \theta_i$ and $l > l_0$, where*

$$l_0 := \max\Big\{\lambda_{\max}\Big(\sqrt{\frac{\theta^2}{4} + \alpha d_{\min}} - \frac{\theta}{2}\Big), \lambda_{\min}\big(\sqrt{\alpha d_{\max}}\big)\Big\} \tag{2.42}$$

(cf. Notation 2.2.1). Then the initial value problems

$$P' = P^2 + \theta P - \alpha d_{\min} I \qquad P(T) = \frac{l}{\lambda_{\max}} I$$

and

$$Q' = Q^2 - \alpha d_{\max} I \qquad Q(T) = \frac{l}{\lambda_{\min}} I$$

possess unique positive definite solutions $P(l, \cdot)$ respectively $Q(l, \cdot)$ on $(-\infty, T]$. P and Q are given by

$$P(l, t) = p(l, t) I, \tag{2.43}$$
$$Q(l, t) = q(l, t) I, \tag{2.44}$$

where

$$p(l, t) := \sqrt{\frac{\theta^2}{4} + \alpha d_{\min}} \coth\Big(\sqrt{\frac{\theta^2}{4} + \alpha d_{\min}}(T - t) + \kappa_1(l)\Big) - \frac{\theta}{2}, \tag{2.45}$$
$$q(l, t) := \sqrt{\alpha d_{\max}} \coth\Big(\sqrt{\alpha d_{\max}}(T - t) + \kappa_2(l)\Big) \tag{2.46}$$

for $\theta + \alpha d_{\min} > 0$ respectively $\alpha d_{\max} > 0$ with

$$\kappa_1(l) := \operatorname{arcoth}\Big(\frac{\frac{l}{\lambda_{\max}} + \frac{\theta}{2}}{\sqrt{\frac{\theta^2}{4} + \alpha d_{\min}}}\Big) > 0,$$
$$\kappa_2(l) := \operatorname{arcoth}\Big(\frac{\frac{l}{\lambda_{\min}}}{\sqrt{\alpha d_{\max}}}\Big) > 0$$

and

$$p(l, t) := \frac{1}{T - t + \frac{\lambda_{\max}}{l}}, \tag{2.47}$$
$$q(l, t) := \frac{1}{T - t + \frac{\lambda_{\min}}{l}} \tag{2.48}$$

for $\theta = \alpha d_{\min} = 0$ *respectively* $\alpha d_{\max} = 0$.

Proof. Let us first consider the scalar initial value problem

$$\begin{aligned} p' &= p^2 + \theta p - \alpha d_{\min} \\ p(T) &= \frac{l}{\lambda_{\max}}. \end{aligned} \tag{2.49}$$

By Section 2.2.2, $p(l,t)$ as in Equation (2.45) respectively Equation (2.47) solves (2.49). We have

$$0 < p(l,t) < \infty \quad \text{for all } t \in (-\infty, T]$$

and

$$\begin{aligned} P'(l,t) &= p'(l,t)I = (p^2(l,t) + \theta p(l,t) - \alpha d_{\min})I = P^2(l,t) + \theta P(l,t) - \alpha d_{\min} I, \\ P(l,T) &= p(l,T)I = \frac{l}{\lambda\text{max}} I \end{aligned}$$

for $P(l,t)$ as in Equation (2.43). By uniqueness of the solution, this establishes the assertions that P as in Equation (2.43) solves the postulated Riccati matrix differential equation uniquely in $(-\infty, T]$ and that $P(l,t) > 0$ for all $t \in (-\infty, T]$.

Setting $\theta = 0$ and replacing $d_{\min}$ and $\lambda_{\max}$ by $d_{\max}$ and $\lambda_{\min}$, establishes Equation (2.44), the existence of $Q(l,\cdot)$ on $(-\infty, T]$, $0 < q(l,t) < \infty$ on $(-\infty, T]$ and thus $Q(l,t) > 0$. $\square$

Lemma 2.2.7 enables us to prove that the Initial Value Problem (2.18) admits a positive definite solution on $(-\infty, T]$. The matrices P and Q turn out to be the desired bounds for $\sqrt{\Lambda^{-1}} C(l,t) \sqrt{\Lambda^{-1}}$. Additionally, we introduce the simpler bounds $\tilde{p}(l,t)I$ and $\tilde{q}(l,t)I$, which simplify the calculations in several proofs of Sections 2.3 and 2.4. Note that all results can be obtained without these simpler bounds by more tedious calculations.

Theorem 2.2.8. *Let* $l > l_0$ *for* l_0 *as in Equation* (2.42)*. The matrix differential equation given by*

$$\begin{aligned} C' &= C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha \Sigma \\ C(T) &= lI \end{aligned} \tag{2.50}$$

possesses a unique solution $C(l,\cdot)$ *on* $(-\infty, T]$*. The solution is symmetric for all* $t \in (-\infty, T]$ *and*

$$\tilde{p}(l,t)I \le P(l,t) \le \sqrt{\Lambda^{-1}} C(l,t) \sqrt{\Lambda^{-1}} \le Q(l,t) \le \tilde{q}(l,t)I, \tag{2.51}$$

where P *and* Q *are as in Lemma 2.2.7, and* $\tilde{p}$ *and* $\tilde{q}$ *are given by*

$$\tilde{p}(l,t) := \frac{1}{T - t + \frac{2\lambda_{\max}}{2l + \theta \lambda_{\max}}} - \frac{\theta}{2},$$

$$\tilde{q}(l,t) := \frac{1}{T - t + \frac{\lambda_{\min}}{l - \lambda_{\min}\sqrt{\alpha d_{\max}}}} + \sqrt{\alpha d_{\max}}.$$

In particular, $C(l,t)$ is positive definite for all $t \in (-\infty, T]$.

Proof. Let $C(l,t)$ be a solution of (2.50) on some interval $(t_1, T]$; note that there exists a local solution by the Picard-Lindelöf theorem. The symmetry of Λ, Σ and the initial value $C(l,T) = lI$ imply that $C(l,t)$ is symmetric on $(t_1, T]$.

Let now $\hat{P} := \Lambda P$ and $\hat{Q} := \Lambda Q$ for P and Q as in Lemma 2.2.7. Then $\hat{P}(l,t)$ solves

$$\begin{aligned} \hat{P}' &= \hat{P}\Lambda^{-1}\hat{P} + \theta\hat{P} - \alpha d_{\min}\Lambda \\ \hat{P}(T) &= \frac{l}{\lambda_{\max}}\Lambda \end{aligned}$$

and $\hat{Q}(l,t)$ solves

$$\begin{aligned} \hat{Q}' &= \hat{Q}\Lambda^{-1}\hat{Q} - \alpha d_{\max}\Lambda \\ \hat{Q}(T) &= \frac{l}{\lambda_{\min}}\Lambda \end{aligned}$$

on $(-\infty, T]$. As $P, Q > 0$ and P and Q commute with Λ, we have

$$\hat{P}(l,t), \hat{Q}(l,t) > 0.$$

Assume that

$$\{t \in (t_1, T] \,|\, C(l,t) \text{ is not positive definite}\} \neq \emptyset \tag{2.52}$$

and define

$$\tau := \sup\{t \in (t_1, T] \,|\, C(l,t) \text{ is not positive definite}\}.$$

As $C(l,T) = lI > 0$ and $C(l,\cdot)$ is continuous, there exists an $\epsilon > 0$ such that $C(l,t) > 0$ for $t \in (T-\epsilon, T]$ and thus $\tau < T$. We apply Theorem 2.2.2 (ii) to $\bar{P} := -\hat{P}$ and $\bar{C} := -C$ on $[\tau, T]$. We have

$$\bar{P}(l,T) = -\frac{l}{\lambda_{\max}}\Lambda \geq -lI = \bar{C}(l,T)$$

and

$$\begin{aligned} \bar{P}' &= -\bar{P}\Lambda^{-1}\bar{P} + \theta\bar{P} + \alpha d_{\min}\Lambda, \\ \bar{C}' &= -\bar{C}\Lambda^{-1}\bar{C} + \bar{C}\bar{\tilde{C}}\bar{C} + \alpha\Sigma \\ &= -\bar{C}\Lambda^{-1}\bar{C} + \theta\bar{C} + \big(\alpha\sqrt{\Lambda}D\sqrt{\Lambda} + \bar{C}\bar{\tilde{C}}\bar{C} - \theta\bar{C}\big). \end{aligned}$$

Let now $x \in \mathbb{R}^n$. Applying Corollary 2.2.5 to $-\bar{C}$, we obtain

$$\begin{aligned} &x^\top\Big(\alpha\sqrt{\Lambda}D\sqrt{\Lambda} + \bar{C}\bar{\tilde{C}}\bar{C} - \theta\bar{C} - \alpha d_{\min}\Lambda\Big)x \\ &\quad = \alpha x^\top(\sqrt{\Lambda}(D - d_{\min}I)\sqrt{\Lambda})x + x^\top(\bar{C}\bar{\tilde{C}}\bar{C} - \theta\bar{C})(t)x \geq 0. \end{aligned}$$

As $\Lambda > 0$, Theorem 2.2.2 (ii) implies

$$\bar{C}(l,t) \le \bar{P}(l,t)$$

and therefore

$$0 < \hat{P}(l,t) \le C(l,t)$$

on $(\tau, T]$. By continuity of $C(l,\cdot)$, we have

$$0 < \hat{P}(l,\tau) \le C(l,\tau)$$

and thus $C(l,t) > 0$ in some neighborhood of τ, a contradiction to Assumption (2.52). Hence, $C(l,t)$ is positive definite on the whole interval $(t_1, T]$. Applying Theorem 2.2.2 (ii) in the same way as above again, yields that we may choose $t_1 = -\infty$ and that

$$0 < \hat{P}(l,t) \le C(l,t)$$

on $(-\infty, T]$.

Let now $\bar{Q} := -\hat{Q}$. We have

$$\bar{Q}(l,T) = -\frac{l}{\lambda_{\min}}\Lambda \le -lI = \bar{C}(l,T)$$

and

$$\begin{aligned}
\bar{Q}' &= -\bar{Q}\Lambda^{-1}\bar{Q} + \alpha d_{\max}\Lambda, \\
\bar{C}' &= -\bar{C}\Lambda^{-1}\bar{C} + (\bar{C}\tilde{\bar{C}}\bar{C} + \alpha\sqrt{\Lambda}D\sqrt{\Lambda}).
\end{aligned}$$

Then for $x \in \mathbb{R}^n$ (note that $\bar{C}\tilde{\bar{C}}\bar{C}(l,t) < 0$ for all t),

$$x^\top(\alpha d_{\max}\Lambda - \bar{C}\tilde{\bar{C}}\bar{C} - \alpha\sqrt{\Lambda}D\sqrt{\Lambda})x = \alpha x^\top\left(\sqrt{\Lambda}(d_{\max}I - D)\sqrt{\Lambda}\right)x - x^\top\bar{C}\tilde{\bar{C}}\bar{C}(l,t)x \ge 0.$$

Thus,

$$\bar{Q}(l,t) \le \bar{C}(l,t)$$

by Theorem 2.2.2 (ii) on $(-\infty, T]$.

Combining the results above, we obtain

$$0 < P(l,t) \le \sqrt{\Lambda^{-1}}C(l,t)\sqrt{\Lambda^{-1}} \le Q(l,t)$$

on $(-\infty, T]$ as Λ commutes with P and Q.

Corollary 2.2.3 directly implies

$$\begin{aligned}
\tilde{p}(l,t)I &\le P(l,t) \quad \text{and} \\
\tilde{q}(l,t)I &\ge Q(l,t).
\end{aligned}$$

$\square$

2.3. Penalizing non-liquidation

In this section we solve the Optimization Problem ($\widetilde{\text{OPT}}$). In Section 2.1.3 we obtained candidates both for the value function and for the optimal strategy (cf. Lemma 2.1.5 and Corollary 2.1.6). In Section 2.2.4 we proved that these candidates are well-defined. We are thus able to state the main result of this section.

Theorem 2.3.1. *Let $l \geq l_0$ for l_0 as in Equation* (2.42) *and let*

$$C(l,t) = (c_{i,j}(l,t))_{i,j=1,\dots,n}$$

be the unique solution of the Initial Value Problem (2.18)*. Then the value function of the Optimization Problem* ($\widetilde{\text{OPT}}$) *is given by*

$$\tilde{v}(l,t,x) = x^\top C(l,t)x \tag{2.53}$$

and the $\mathbb{P} \otimes \lambda$ - almost surely unique optimal strategy is given by $u^*(l) := (\xi^*(l), \eta^*(l))$,

$$\xi^*(l) := \xi^*(l,t,x) := \Lambda^{-1}C(l,t)x, \tag{2.54}$$

$$\eta^*(l) := \eta^*(l,t,x) := \tilde{I}\bar{C}(l,t)C(l,t)x, \tag{2.55}$$

where $\tilde{I} = (e_{i,j})_{i,j=1,\dots,n}$ is the diagonal matrix with

$$e_{i,i} = \begin{cases} 1 & \text{if } \theta_i > 0 \\ 0 & \text{else} \end{cases}$$

and

$$\bar{C}(l,t) := \operatorname{diag}\Big(\frac{1}{c_{i,i}(l,t)}\Big).$$

We use the following notation.

Notation 2.3.2. *(i) The independent Poisson processes π_i, $i = 1, \dots n$, jump at different times almost surely, and the number of jumps in $[t,T]$ is almost surely finite. We denote the jump times of π by $(\tau_j)_{j \in \mathbb{N}}$, where $\tau_j < \tau_{j+1}$ ($j \in \mathbb{N}$) almost surely.*

(ii) Given the Markovian control $u^*(l)$, the Stochastic Differential Equation (2.1) *possesses a unique solution. We denote the process controlled by $u^*(l)$ by*

$$X^*(l,s) := X^{u^*(l)}(s).$$

For the proof of Theorem 2.3.1 we have to show that $u^*(l)$ is an admissible trading strategy first. This is accomplished in Section 2.3.1 by using the bounds of C obtained in Section 2.2.4. Section 2.3.2 finishes the proof of Theorem 2.3.1.

2.3.1. Admissibility of the candidate optimal strategy

In order to prove admissibility of $u^*(l)$, we require an upper bound for the process $\|X^*(l,s)\|_2$. Using Gronwall's inequality pathwise inductively on the time-intervals

$$[\tau_i \wedge T, \tau_{i+1} \wedge T)$$

and interlacing the jumps, this can be achieved by applying the upper and lower bounds of $C(l,s)$ from Section 2.2.4.

Lemma 2.3.3. *Let $l > l_0$ for l_0 as in Equation* (2.42)*, $t \in [0,T)$ and $x \in \mathbb{R}^n$ be the portfolio position at time t. Then for $s \in [t,T)$,*

$$\begin{aligned}
&X^*(l,s)^\top \Lambda X^*(l,s) \\
&\quad \le x^\top \Lambda x \exp\Big(-2\int_t^s p(l,u)du\Big) \prod_{t\le\tau_i\le s} (1+n^2)\frac{q(l,\tau_i)}{p(l,\tau_i)} && (2.56)\\
&\quad \le x^\top \Lambda x \exp\left(\theta(s-t)\right) \frac{\left(T-s+\frac{2\lambda_{\max}}{2l+\theta\lambda_{\max}}\right)^2}{\left(T-t+\frac{2\lambda_{\max}}{2l+\theta\lambda_{\max}}\right)^2} \prod_{t\le\tau_i\le s} (1+n^2)\frac{q(l,\tau_i)}{p(l,\tau_i)} \quad a.s., && (2.57)
\end{aligned}$$

where p and q are as in Equations (2.45) *and* (2.46) *respectively* (2.47) *and* (2.48)*.*

Proof. Let $i \in \mathbb{N}$. On $\{\tau_i < T\}$, $X^*(l,\cdot)$ solves the initial value problem

$$\begin{aligned}
X' &= -\Lambda^{-1}C(l)X \\
X(\tau_i) &= X^*(l,\tau_i)
\end{aligned}$$

for $s \in [\tau_i, \tau_{i+1} \wedge T)$, and therefore

$$Y(l,\cdot) := \sqrt{\Lambda}X^*(l,\cdot)$$

solves the initial value problem

$$\begin{aligned}
Y' &= -\sqrt{\Lambda^{-1}}C(l)\sqrt{\Lambda^{-1}}Y \\
Y(\tau_i) &= \sqrt{\Lambda}X^*(l,\tau_i)
\end{aligned}$$

for $s \in [\tau_i, \tau_{i+1} \wedge T)$. We set

$$Z(l,s) := Y(l,s)^\top Y(l,s) = X^*(l,s)^\top \Lambda X^*(l,s) \tag{2.58}$$

and obtain

$$\begin{aligned}
Z'(l,s) &= Y'(l,s)^\top Y(l,s) + Y(l,s)^\top Y'(l,s) \\
&= -2Y(l,s)^\top \left(\sqrt{\Lambda^{-1}}C(l,s)\sqrt{\Lambda^{-1}}\right) Y(l,s) \\
&\le -2p(l,s)Y(l,s)^\top Y(l,s)
\end{aligned}$$

$$= -2p(l,s)Z(l,s)$$

by Theorem 2.2.8. Gronwall's inequality thus implies that

$$Z(l,s) \le Z(l,\tau_i)\exp\Big(-2\int_{\tau_i}^{s} p(l,u)du\Big)$$

and

$$Z(l,(\tau_{i+1}\wedge T)-) \le Z(l,\tau_i)\exp\Big(-2\int_{\tau_i}^{\tau_{i+1}\wedge T} p(l,u)du\Big). \tag{2.59}$$

Let now $s \in [t,T)$. By Corollary 2.2.6,

$$0 < C(l,s) \le n\bar{C}(l,s)^{-1}$$

and therefore

$$C\bar{C}C\bar{C}C(l,s) \le nC\bar{C}C(l,s) \le n^2 C(l,s), \tag{2.60}$$

where the last inequality follows from Corollary 2.2.6 again. On $\{\tau_{i+1} < T\}$, this implies

$$\begin{aligned}
&X^*(l,\tau_{i+1})^\top C(l,\tau_{i+1})X^*(l,\tau_{i+1})\\
&\quad = \big(X^*(l,\tau_{i+1}-) - \eta^*(l,\tau_{i+1},X^*(l,\tau_{i+1}-))\big)^\top\\
&\qquad\qquad C(l,\tau_{i+1})\big(X^*(l,\tau_{i+1}-) - \eta^*(l,\tau_{i+1},X^*(l,\tau_{i+1}-))\big)\\
&\quad = X^*(l,\tau_{i+1}-)^\top C(l,\tau_{i+1})X^*(l,\tau_{i+1}-)\\
&\qquad\qquad - 2X^*(l,\tau_{i+1}-)^\top \tilde{I}C\bar{C}C(l,\tau_{i+1})X^*(l,\tau_{i+1}-)\\
&\qquad\qquad\qquad + X^*(l,\tau_{i+1}-)^\top \tilde{I}C\bar{C}C\bar{C}C(l,\tau_{i+1})X^*(l,\tau_{i+1}-)\\
&\quad \overset{(2.60)}{\le} (1+n^2)X^*(l,\tau_{i+1}-)^\top C(l,\tau_{i+1})X^*(l,\tau_{i+1}-)
\end{aligned} \tag{2.61}$$

and thus

$$\begin{aligned}
&Y(l,\tau_{i+1})^\top Y(l,\tau_{i+1})\\
&\quad \le Y(l,\tau_{i+1})^\top \left(\frac{\sqrt{\Lambda^{-1}}C(l,\tau_{i+1})\sqrt{\Lambda^{-1}}}{p(l,\tau_{i+1})}\right) Y(l,\tau_{i+1})\\
&\quad = \frac{1}{p(l,\tau_{i+1})}X^*(l,\tau_{i+1})^\top C(l,\tau_{i+1})X^*(l,\tau_{i+1})\\
&\quad \overset{(2.61)}{\le} \frac{1+n^2}{p(l,\tau_{i+1})}X^*(l,\tau_{i+1}-)^\top C(l,\tau_{i+1})X^*(l,\tau_{i+1}-)\\
&\quad = \frac{1+n^2}{p(l,\tau_{i+1})}Y(l,\tau_{i+1}-)^\top \left(\sqrt{\Lambda^{-1}}C(l,\tau_{i+1})\sqrt{\Lambda^{-1}}\right) Y(l,\tau_{i+1}-)\\
&\quad \le (1+n^2)\frac{q(l,\tau_{i+1})}{p(l,\tau_{i+1})}Y(l,\tau_{i+1}-)^\top Y(l,\tau_{i+1}-) \quad \text{a.s.}
\end{aligned} \tag{2.62}$$

Using Inequalities (2.59) and (2.62), we obtain Inequality (2.56) inductively. Inequality (2.57) follows from

$$\tilde{p}(l,u) = \frac{1}{T-u+\frac{2\lambda_{\max}}{2l+\theta\lambda_{\max}}} - \frac{\theta}{2} \leq p(l,u)$$

(cf. Inequality (2.51)). □

The bound obtained in Lemma 2.3.3 enables us to prove that $u^*(l)$ fulfills the moment conditions of Definition 2.1.1 (ii) and that it is an admissible trading strategy.

Proposition 2.3.4. *Let $l > l_0$ for l_0 as in Equation* (2.42) *and $(t,x) \in [0,T)\times\mathbb{R}^n$. Then $u^*(l) = (\xi^*(l), \eta^*(l))$ for $\xi^*(l)$ and $\eta^*(l)$ as in Equations* (2.54) *and* (2.55) *respectively is admissible.*

Proof. Definition 2.1.1 (i) and (iii) are clearly satisfied. By Lemma 2.3.3, we obtain for $k \geq 1$,

$$\begin{aligned}\|X^*(l,s)\|_2^k &\leq \Big(\frac{1}{\lambda_{\min}} X^*(l,s)^\top \Lambda X^*(l,s)\Big)^{\frac{k}{2}} \\ &\leq \Big(\frac{1}{\lambda_{\min}} x^\top \Lambda x\Big)^{\frac{k}{2}} \exp\big(\tfrac{k}{2}\theta(s-t)\big) \frac{\big(T-s+\frac{2\lambda_{\max}}{2l+\theta\lambda_{\max}}\big)^k}{\big(T-t+\frac{2\lambda_{\max}}{2l+\theta\lambda_{\max}}\big)^k} \Big(\prod_{t\leq\tau_i\leq s}(1+n^2)\frac{q(l,\tau_i)}{p(l,\tau_i)}\Big)^{\frac{k}{2}}.\end{aligned}$$

Thus, as $0 < \frac{q(l,s)}{p(l,s)}$ is bounded on $[0,T]$, there exist constants $\tilde{K} = \tilde{K}(k)$ and $\bar{K} = \bar{K}(k)$ independent of s such that

$$\mathbb{E}_{t,x}\big[\|X^*(l,s)\|_2^k\big] \leq \tilde{K}\mathbb{E}_{t,x}\Big[\bar{K}^{\#\{i|t\leq\tau_i\leq T\}}\Big] = \tilde{K}\mathbb{E}_{t,x}\Big[\bar{K}^{\tilde{\pi}(T-t)}\Big] = \tilde{K}\exp(\theta\bar{K}(T-t)),$$

where $\tilde{\pi}$ is a Poisson process with intensity $\theta = \sum_i \theta_i$.

Let now $\|\cdot\|_{2,2}$ denote the matrix norm induced by the space $(\mathbb{R}^n, \|\cdot\|_2)$. Note that $\|\cdot\|_{2,2}$ is the spectral norm on $\mathbb{R}^{n\times n}$ and therefore (see, e.g., Bernstein [2005], Theorem 8.4.9)

$$\|A\|_{2,2} \leq \|B\|_{2,2} \quad \text{for} \quad 0 \leq A \leq B.$$

Using Theorem 2.2.8, we deduce

$$\begin{aligned}\mathbb{E}_{t,x}\Big[\int_t^T \|\xi^*(l,s,X^*(s))\|_2^4 ds\Big] &\leq \mathbb{E}_{t,x}\Big[\int_t^T \|\Lambda^{-1}\|_{2,2}^4 \|C(l,s)\|_{2,2}^4 \|X^*(l,s)\|_2^4 ds\Big] \\ &\leq \mathbb{E}_{t,x}\Big[\int_t^T q(l,s)^4 \tilde{K}\exp(\bar{K}(T-t))ds\Big] < \infty\end{aligned}$$

and

$$\mathbb{E}_{t,x}\Big[\int_t^T \|\eta^*(l,s,X^*(s-))\|_2^8 ds\Big] \le \mathbb{E}_{t,x}\Big[\int_t^T \|\underbrace{\bar{C}(l,s)}_{\le nC(l,s)^{-1} \text{ by Corollary 2.2.6}}\|_{2,2}^8 \|C(l,s)\|_{2,2}^8 \|X^*(l,s-)\|_2^8 ds\Big]$$
$$\le \mathbb{E}_{t,x}\Big[\int_t^T \frac{n^8 \lambda_{max}^8 q(l,s)^8}{\lambda_{\min}^8 p(l,s)^8} \tilde{K} \exp(\bar{K}(T-t)) ds\Big] < \infty.$$

□

2.3.2. Proof of Theorem 2.3.1

We are now ready do prove the central result of this section. The proof is a verification argument using Itô's formula.

Proof of Theorem 2.3.1. Let $l > l_0$, $(t,x) \in [0,T) \times \mathbb{R}^n$ and $u = (\xi, \eta) \in \tilde{\mathbb{A}}(t,x)$. We apply Itô's formula (see, e.g., the book by Øksendal and Sulem [2007]) to the function

$$w(l,t,X^u(t)) = X^u(t)^\top C(l,t) X^u(t).$$

$$\begin{aligned}
&w(l,t,x)\\
&\quad = w(l,T,X^u(T)) + \int_t^T \nabla_x w(l,s,X^u(s))\xi(s) - \frac{\partial w}{\partial s}(l,s,X^u(s))ds\\
&\qquad + \int_t^T \Big(\sum_{i=1}^n w(l,s,X^u(s-)) - w\big(l,s,X^u(s-) - \eta_i(s)e_i\big)\Big)\pi_i(ds)\\
&\quad \le w(l,T,X^u(T)) + \int_t^T f(X^u(s),\xi(s))ds\\
&\qquad + \int_t^T \Big(\sum_{i=1}^n w(l,s,X^u(s-)) - w\big(l,s,X^u(s-) - \eta_i(s)e_i\big)\Big)\pi_i(ds)\\
&\qquad - \int_t^T \Big(\sum_{i=1}^n \theta_i\big(w(l,s,X^u(s-)) - w\big(l,s,X^u(s-) - \eta_i(s)e_i\big)\Big)ds \qquad (2.63)\\
&\quad = w(l,T,X^u(T)) + \int_t^T f(X^u(s),\xi(s))ds\\
&\qquad + \sum_{i=1}^n \int_t^T \Big(w(l,s,X^u(s-)) - w\big(l,s,X^u(s-) - \eta_i(s)e_i\big)\Big)M_i(ds),
\end{aligned}$$

where M_i is the compensated Poisson process $M_i(s) := \pi_i(s) - \theta_i s$ and Inequality (2.63) follows from Corollary 2.1.6. Furthermore, we have (pathwise) equality in (2.63) if and only if $u = u^*$ λ - a.s.

Taking expectations on both sides, we obtain

$$\begin{aligned} w(l,t,x) &\leq \tilde{J}(l,t,x,u) \\ &\qquad + \sum_{i=1}^{n} \mathbb{E}_{t,x}\Big[\int_t^T \Big(w(l,s,X^u(s-)) - w\big(l,s,X^u(s-) - \eta_i(s)e_i\big)\Big) M_i(ds)\Big], \end{aligned} \tag{2.64}$$

with equality if and only if $u = u^*$ $\mathbb{P} \otimes \lambda$ - a.s.

It remains to show that the stochastic integrals in Inequality (2.64) are martingales. To this end, we compute

$$\begin{aligned} &\mathbb{E}_{t,x}\Big[\int_t^T |w(l,s,X^u(s-)) - w\big(l,s,X^u(s-) - \eta_i(s)e_i\big)|^2 ds\Big] \\ &= \mathbb{E}_{t,x}\Big[\int_t^T |2X^u(s-)^\top C(l,s)e_i\eta_i(s) - \eta_i(s)^2 c_{i,i}(s,l)|^2 ds\Big] \\ &\leq \mathbb{E}_{t,x}\Big[\int_t^T 2|2X^u(s-)^\top C(l,s)e_i\eta_i(s)|^2 ds\Big] + \mathbb{E}_{t,x}\Big[\int_t^T 2|\eta_i(s)^2 c_{i,i}(l,s)|^2 ds\Big] \\ &\leq 8\mathbb{E}_{t,x}\Big[\int_t^T \|X^u(s-)^\top\|_2^2 \|C(l,s)\|_{2,2}^2 |\eta_i(s)|^2 ds\Big] + 2\mathbb{E}_{t,x}\Big[\int_t^T |\eta_i(s)|^4 |c_{i,i}(l,s)|^2 ds\Big] \\ &\leq 8\lambda_{\max}^2\big(\max_{s\in[t,T]} q(l,s)^2\big)\mathbb{E}_{t,x}\Big[\int_t^T \|X^u(s-)\|_2^4 ds\Big]^{\frac{1}{2}} \mathbb{E}_{t,x}\Big[\int_t^T \|\eta(s)\|_2^4 ds\Big]^{\frac{1}{2}} \\ &\qquad + 2\lambda_{\max}^2\big(\max_{s\in[t,T]} q(l,s)^2\big)\mathbb{E}_{t,x}\Big[\int_t^T \|\eta(s)\|_2^4 ds\Big] \qquad (2.65) \\ &< \infty \end{aligned}$$

by Definition 2.1.1 (ii) and Proposition 2.1.4, where Inequality (2.65) follows from Hölder's inequality. As $\langle M_i\rangle(s) = \theta_i s$, this finishes the proof. □

2.4. Optimal liquidation

We finally solve the Optimization Problem (OPT) in this section. A candidate for the solution is the limit of the solution of the Optimization Problem ($\widetilde{\text{OPT}}$) for $l \to \infty$. In

Section 2.4.1 we prove that the element-wise limit

$$C(t) := \lim_{l\to\infty} C(l,t)$$

exists for $t \in [0,T)$ so that the candidate for the value function of (OPT),

$$w(t,x) := x^\top C(t)x,$$

is well-defined. We also obtain upper and lower bounds for C and deduce that the limit

$$u^* := \lim_{l\to\infty} u^*(l)$$

exists as well. In Section 2.4.2 we analyze the candidate optimal strategy u^* and the corresponding controlled process

$$X^* := X^{u^*}.$$

In particular, we show that u^* is an admissible liquidation strategy. Using Sections 2.4.1 and 2.4.2, we verify that $w(t,x)$ is indeed the value function and that u^* is the optimal strategy in Section 2.4.3.

2.4.1. The candidate value function

We start by computing the limits of the functions $p(l)$ and $q(l)$ given by Equations (2.45) and (2.46) respectively (2.47) and (2.48) and obtain candidates for an upper and a lower bound of $\lim\limits_{l\to\infty} C(l,t)$.

Lemma 2.4.1. *Let $t \in [0,T)$ and $p(l)$ and $q(l)$ as in Equations* (2.45) *and* (2.46) *respectively* (2.47) *and* (2.48)*. Then*

$$\lim_{l\to\infty} p(l,\cdot) = p(\cdot), \quad \lim_{l\to\infty} q(l,\cdot) = q(\cdot)$$

compactly and strictly increasing on $[0,T)$, where p and q are given by

$$p(t) := \sqrt{\frac{\theta^2}{4} + \alpha d_{\min}} \coth\left(\sqrt{\frac{\theta^2}{4} + \alpha d_{\min}}(T-t)\right) - \frac{\theta}{2}, \tag{2.66}$$

$$q(t) := \sqrt{\alpha d_{\max}} \coth\left(\sqrt{\alpha d_{\max}}(T-t)\right) \tag{2.67}$$

for $\theta + \alpha d_{\min} > 0$ respectively $\alpha d_{\max} > 0$ and

$$p(t) := \frac{1}{T-t}, \tag{2.68}$$

$$q(t) := \frac{1}{T-t} \tag{2.69}$$

for $\theta = \alpha d_{\min} = 0$ *respectively* $\alpha d_{\max} = 0$. *Furthermore,*

$$p(l,T), q(l,T) \nearrow \infty \qquad \text{as } l \to \infty.$$

Proof. Point-wise convergence and the formulae for the limits are straightforward by Equations (2.45) and (2.46) respectively by Equations (2.47) and (2.48). Strict monotonicity follows from the fact that the initial values are strictly increasing in l. Finally, compact convergence follows from these observations by Dini's theorem (see, e.g., Courant and Hilbert [1953]). $\square$

In the following Lemma we prove monotonicity of the sequence $(C(l,t))_l$ in the sense of Notation 2.2.1 (i). Combining this with the candidate bounds obtained in Lemma 2.4.1, we can prove the existence of

$$\lim_{l\to\infty} C(l,t)$$

in Proposition 2.4.3 below.

Lemma 2.4.2. *For fixed* $t \in (-\infty, T]$, $C(l,t)$ *is strictly increasing in* l *on* (l_0,∞) *for* l_0 *as in Equation* (2.42).

Proof. We differentiate C with respect to l (by differentiating with respect to initial values), denote the partial derivative by

$$C_l := (c^l_{i,j})_{i,j=1,\dots n} := \frac{\partial C}{\partial l}$$

and show that $C_l(l,t) > 0$ on $(-\infty, T]$ (Formula (2.71) below).

Note that we can exchange partial derivatives and obtain the initial value problem

$$\begin{aligned} \frac{\partial}{\partial t} C_l(l,t) &= \frac{\partial}{\partial l}\frac{\partial}{\partial t} C(l,t) = \frac{\partial}{\partial l}\Big(C\Lambda^{-1}C + C\tilde{C}C - \alpha\Sigma\Big)(l,t) \\ &= \Big(C_l A + A^\top C_l - B^\top \operatorname{diag}(c^l_{i,i}) B\Big)(l,t) \\ C_l(l,T) &= I, \end{aligned} \tag{2.70}$$

where

$$A(l,t) := (\Lambda^{-1} + \tilde{C}(l,t))C(l,t), \quad B(l,t) := \operatorname{diag}\Big(\frac{\sqrt{\theta_i}}{c_{i,i}(l,t)}\Big)C(l,t).$$

The Initial Value Problem (2.70) is a linear matrix differential equation for C_l ($l > 0$ fixed) and is thus solvable on the whole interval $(-\infty, T]$ (with solution C_l).

We consider the linear matrix differential equation

$$x' = -Ax$$

and its fundamental matrix Φ with initial condition $\Phi(l,T) = I$, i.e.,

$$\Phi' = -A\Phi, \quad \Phi(l,T) = I$$

and $\Phi(l,t)$ is regular for all $t \in (-\infty, T]$. We define

$$Y(l,t) := \Phi(l,t)^\top C_l(l,t)\Phi(l,t),$$

thus

$$C_l(l,t) = (\Phi(l,t)^{-1})^\top Y(l,t)\Phi(l,t)^{-1}$$

and

$$\begin{aligned} Y' &= -\Phi^\top A^\top C_l \Phi + \Phi^\top (C_l A + A^\top C_l - B^\top \operatorname{diag}(c^l_{i,i}) B)\Phi - \Phi^\top C_l A \Phi \\ &= -\Phi^\top B^\top \operatorname{diag}(c^l_{i,i}) B\Phi, \\ &= -\Phi^\top B^\top \operatorname{diag}((\Phi^{-1})^\top Y(l,t)\Phi^{-1}) B\Phi, \\ Y(l,T) &= I. \end{aligned}$$

This is a linear matrix differential equation for Y and has a symmetric solution on $(-\infty, T]$. In the following we show that $Y(l,t) \geq I > 0$ for all $t \in (-\infty, T]$. To this end, define

$$\tau := \inf\{t | Y(l,\cdot) > 0 \text{ on } [t,T]\} < T.$$

Then $\Phi^\top B^\top \operatorname{diag}((\Phi^{-1})^\top Y(l,t)\Phi^{-1}) B\Phi \geq 0$ and hence $Y'(l,t) \leq 0$ on $(\tau, T]$. Thus,

$$Y(l,t) = Y(l,T) - \int_t^T Y'(l,s)ds \geq Y(l,T) = I$$

on $(\tau, T]$. Assume that $\tau > -\infty$. Continuity of $Y(l,\cdot)$ yields $Y(l,\tau) \geq I > 0$ and $Y(l,t) > 0$ on the interval $[\tau - \epsilon, \tau + \epsilon]$ for some $\epsilon > 0$, a contradiction. Therefore for all $t \in (-\infty, T]$,

$$C_l(l,t) = (\Phi(l,t)^{-1})^\top Y(l,t)\Phi(l,t)^{-1} > 0, \tag{2.71}$$

finishing the proof. □

We are now ready to prove the existence of the limit $\lim_{l\to\infty} C(l,t)$ by combining the two preceding lemmata.

Proposition 2.4.3. *The element-wise limit of the value function matrix*

$$C(t) := \lim_{l\to\infty} C(l,t)$$

exists on $[0,T)$*, and* $C(l,\cdot)$ *converges compactly to* C *on* $[0,T)$*. Furthermore,*

$$\lim_{l\to\infty} c_{min}(l,T) = \infty.$$

Proof. The existence of the element-wise limit of $(C(l,t))_{l>0}$ follows directly from the monotonicity (cf. Lemma 2.4.2) and the boundedness by $\Lambda q(t) I$ for $q(t)$ as in Equation (2.67) respectively (2.69).

Compact convergence follows by Dini's theorem due to the monotonicity. □

We can deduce upper and lower bounds for the matrix C and that C solves the differential equation

$$C' = C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha\Sigma.$$

This result is essential for the verification argument in Section 2.4.3.

Theorem 2.4.4. *For $t \in [0, T)$,*

$$0 < P(t) \leq \sqrt{\Lambda^{-1}} C(t) \sqrt{\Lambda^{-1}} \leq Q(t), \tag{2.72}$$

where

$$P(t) = p(t)I, \quad Q(t) = q(t)I$$

for p and q as in Equations (2.66) *and* (2.67) *respectively* (2.68) *and* (2.69)*; moreover, C solves the matrix differential equation*

$$C' = C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha\Sigma \tag{2.73}$$

on $[0, T)$.

Proof. The inequalities in (2.72) follow directly from the previous results. Furthermore, the compact convergence of $C(l,t)$ (and $\tilde{C}(l,t)$) on $[0,T)$ and the fact that $C(l,\cdot)$ solves the Matrix Differential Equation (2.73) implies that $C'(l,t)$ converges compactly on $[0,T)$ to some matrix D such that $C'(t) = D(t)$ on $[0,T)$, finishing the proof. □

Remark 2.4.5. *For Riccati matrix differential equations, there exists a unique solution Q with*

$$\lim_{s \to T-} q_{\min}(s) = \infty.$$

This solution is called the principal solution *(see, e.g., Coppel [1971]). In this spirit, C is the principal solution of the Matrix Differential Equation* (2.73)*. Note however that it is not entirely clear that C is the only solution of* (2.73) *satisfying*

$$\lim_{s \to T-} c_{\min}(s) = \infty$$

since (2.73) *is* not *a Riccati matrix differential equation.*

2.4.2. The candidate optimal trading strategy and trajectory

By Proposition 2.4.3, we also obtain the existence of the limits of the optimal strategy (cf. the uniform bounds obtained in Lemma 2.4.1):

$$\xi^* := \xi^*(t,x) := \lim_{l \to \infty} \xi^*(l,t,x) = \Lambda^{-1} C(t) x, \tag{2.74}$$

$$\eta^* := \eta^*(t,x) := \lim_{l \to \infty} \eta^*(l,t,x) = \tilde{I}\bar{C}(t) C(t) x. \tag{2.75}$$

It turns out that $u^* := (\xi^*, \eta^*)$ solves the Optimization Problem (OPT). As a first step it is thus necessary to prove that u^* is an admissible liquidation strategy. To this end, we first analyze the controlled process

$$X^*(t) := X^{u^*}(t)$$

and show that

$$\lim_{s \to T-} X^*(s) = 0 \quad \text{a.s.}$$

Proposition 2.4.6. *Let $t \in [0,T)$ and $x \in \mathbb{R}^n$ be the portfolio position at time t.*

(i)

$$X^*(l,\cdot) \overset{l\to\infty}{\longrightarrow} X^*(\cdot) \quad \textit{a.s. compactly on } [t,T).$$

(ii)

$$l \cdot \|X^*(l,T)\|_2^2 \overset{l\to\infty}{\longrightarrow} 0 \quad \textit{a.s. and in } L^1.$$

In particular,

$$X^*(l,T) \overset{l\to\infty}{\longrightarrow} X^*(T) = \lim_{s\to T-} X^*(s) = 0 \quad \textit{a.s.}$$

For proving Proposition 2.4.6, we require the following lemma (cf. also Lemma 2.3.3).

Lemma 2.4.7. *Let $t \in [0,T]$ and $x \in \mathbb{R}^n$ be the portfolio position at time t. Then for $s \in [t,T]$,*

$$X^*(s)^\top \Lambda X^*(s) \leq x^\top \Lambda x \exp\Big(-2 \int_t^s p(u)du\Big) \prod_{t \leq \tau_i \leq s} (1+n^2)\frac{q(\tau_i)}{p(\tau_i)} \tag{2.76}$$

$$\leq x^\top \Lambda x \exp\left(\theta(s-t)\right)\frac{(T-s)^2}{(T-t)^2} \prod_{t \leq \tau_i \leq s} (1+n^2)\frac{q(\tau_i)}{p(\tau_i)}. \tag{2.77}$$

Proof. Inequality (2.76) follows by the same argument as the respective bound for $X^*(l,s)^\top \Lambda X^*(l,s)$ in Lemma 2.3.3. Inequality (2.77) follows from the fact that

$$p(u) \geq \tilde{p}(u) := \lim_{l\to\infty} \tilde{p}(l,u) = \frac{1}{T-u} - \frac{\theta}{2}.$$

□

Proof of Proposition 2.4.6. (i) The spectral norm $\|\cdot\|_{2,2}$ is equivalent to the matrix maximum norm, and therefore the element-wise convergence results from Proposition 2.4.3 transfer to corresponding results for the spectral norm.

Let $t \leq T' < T$. On $\{\tau_i < T'\}$, X^* and $X^*(l)$ solve the respective ordinary differential equations

$$X' = -\xi^*(\cdot, X) = -\Lambda^{-1}CX,$$

$$X' = -\xi^*(l,\cdot,X) = -\Lambda^{-1}C(l)X$$

on the interval $[\tau_i, \tau_{i+1} \wedge T')$. We prove that the assertion follows from the continuous dependence of solutions of ordinary differential equations on the right hand side and initial values. To this end, we first require some preliminary observations.

For $s \in [t, T']$ and $x, y \in \mathbb{R}^n$, we have

$$\|\Lambda^{-1}C(s)x - \Lambda^{-1}C(s)y\|_2 \leq \max_{s\in[t,T']} \|Q(s)\|_{2,2}\|x-y\|_2 =: L\|x-y\|_2$$

(cf. Theorem 2.4.4), i.e., for all $s \in [t, T']$, $\xi^*(s,\cdot)$ is Lipschitz-continuous on $\mathbb{R}^n$ with Lipschitz constant $L = L(T')$ independent of s.

Furthermore, there exits a constant $K_1 \geq 1$ and a random variable $K_2 \geq 1$ such that for $s \in [t, T']$,

$$\|\bar{C}(s)C(s)\|_{2,2} \leq K_1$$

and

$$\begin{aligned}
\|X^*(l,s)\|_2^2 &\leq \frac{1}{\lambda_{\min}} X^*(l,s)^\top \Lambda X^*(l,s) \\
&\leq \frac{1}{\lambda_{\min}} x^\top \Lambda x \exp\Big(-2\int_t^s p(l,u)du\Big) \prod_{t\leq\tau_i\leq T'} (1+n^2)\frac{q(l,\tau_i)}{p(l,\tau_i)} \qquad (2.78)\\
&\leq \frac{1}{\lambda_{\min}} x^\top \Lambda x \prod_{t\leq\tau_i\leq T'} (1+n^2)\frac{q(l,\tau_i)}{p(l,\tau_i)} \\
&\leq K_2,
\end{aligned}$$

where Inequality (2.78) follows from Lemma 2.3.3. Note now that

$$\lim_{t\to T-} \frac{q(t)}{p(t)} = 1,$$

and thus $\frac{q(t)}{p(t)}$ admits a continuous extension to $[0,T]$. Therefore, there exists a constant K_3 such that for all $i \in \mathbb{N}$,

$$\frac{q(\tau_i \wedge T)}{p(\tau_i \wedge T)} \leq K_3$$

and hence

$$\mathbb{E}_{t,x}\Big[\prod_{0\leq\tau_i\leq T} \frac{q(\tau_i)}{p(\tau_i)}\Big] \leq \mathbb{E}_{t,x}\Big[K_3^{\#\{i|0\leq\tau_i\leq T\}}\Big] = \sum_{i=0}^{\infty} K_3^i \frac{\exp(\theta T)(\theta T)^i}{i!} = \exp(K_3) < \infty.$$

Thus K_2 is almost surely finite.

We now set $\tau_0 = t$ and show by induction on $i \in \mathbb{N}$ that for all $\epsilon > 0$, there exits an $l_i > l_0$ such that $l_i \geq l_{i-1}$ and for all $l \geq l_i$, $s \in [t, \tau_i \wedge T']$,

$$\|X^*(l,s) - X^*(s)\|_2 < \epsilon.$$

The assertion is clear for $i = 0$. Let $i > 0$ and $\epsilon > 0$. By the induction hypothesis, there exists $l_{i-1} > l_0$ such that for $l > l_{i-1}$,

$$\|X^*(l,\tau_{i-1}) - X^*(\tau_{i-1})\|_2 < \epsilon \frac{e^{-L(T'-t)}}{6K_1}. \tag{2.79}$$

Note that on $\{\tau_{i-1} \geq T'\}$ the induction step is trivial. We therefore fix some

$$\omega \in \{\tau_{i-1} < T'\}.$$

Let now $l_i \geq l_{i-1}$ such that for $l > l_i$, $s \leq \tau_i \wedge T'$ (recall the uniform convergence of $(C(l,s))_l$ on $[t,T']$, Proposition 2.4.3)

$$\|\Lambda^{-1}C(l,s) - \Lambda^{-1}C(s)\|_{2,2} \leq \|\Lambda^{-1}\|_{2,2}\|C(l,s) - C(s)\|_{2,2} < \epsilon \frac{e^{-L(T'-t)}}{6(T'-t)K_1} \tag{2.80}$$

and

$$\|\bar{C}(l,s)C(l,s) - \bar{C}(s)C(s)\|_{2,2} < \frac{\epsilon}{3K_2(\omega)}. \tag{2.81}$$

By the continuous dependence of solutions of systems of ordinary differential equations on the right hand side and initial values, we have for $s \in [\tau_{i-1}, \tau_i \wedge T')$ (by Inequalities (2.79), and (2.80)),

$$\begin{aligned}\|X^*(l,s,\omega) - X^*(s,\omega)\|_2 &\leq \left(\epsilon \frac{e^{-L(T'-t)}}{6K_1} + (T'-t)\epsilon \frac{e^{-L(T'-t)}}{6(T'-t)K_1}\right) e^{L(T'-t)} \\ &= \frac{\epsilon}{3K_1},\end{aligned}$$

in particular

$$\|X^*(l,(\tau_i \wedge T')-) - X^*((\tau_i \wedge T')-)\|_2 \leq \frac{\epsilon}{3K_1}. \tag{2.82}$$

We can conclude by using the Inequalities (2.81) and (2.82):

$$\begin{aligned}&\|X^*(l,\tau_i(\omega) \wedge T',\omega) - X^*(\tau_i \wedge T',\omega)\|_2 \\ &\quad = \|X^*(l,(\tau_i(\omega) \wedge T')-,\omega) \\ &\qquad - \bar{C}(l,\tau_i(\omega) \wedge T')C(l,\tau_i(\omega) \wedge T')X^*(l,(\tau_i(\omega) \wedge T')-,\omega) \\ &\qquad - X^*((\tau_i(\omega) \wedge T')-,\omega) \\ &\qquad + \bar{C}(\tau_i(\omega) \wedge T')C(\tau_i(\omega) \wedge T')X^*((\tau_i(\omega) \wedge T')-,\omega)\|_2\end{aligned}$$

$$
\begin{aligned}
&\leq \underbrace{\left\|X^*(l,(\tau_i(\omega)\wedge T')-,\omega)-X^*((\tau_i(\omega)\wedge T')-,\omega)\right\|_2}_{\leq\epsilon/(3K_1)\leq\epsilon/3 \text{ by Inequality (2.82)}}\\
&\qquad+\|X^*(l,(\tau_i(\omega)\wedge T')-,\omega)\|_2\\
&\qquad\quad\underbrace{\cdot\|\bar{C}(l,\tau_i(\omega)\wedge T')C(l,\tau_i(\omega)\wedge T')-\bar{C}(\tau_i(\omega)\wedge T')C(\tau_i(\omega)\wedge T')\|_{2,2}}_{\leq\epsilon/3 \text{ by Inequality (2.81)}}\\
&\qquad+\|\bar{C}(\tau_i(\omega)\wedge T')C(\tau_i(\omega)\wedge T')\|_2\\
&\qquad\quad\underbrace{\cdot\|X^*(l,(\tau_i(\omega)\wedge T')-,\omega)-X^*((\tau_i(\omega)\wedge T')-,\omega)\|_2}_{<\epsilon/3 \text{ by Inequality (2.82)}}\\
&<\epsilon
\end{aligned}
$$

as required.

(ii) For fixed $\omega \in \Omega$, we have by Lemma 2.3.3 that

$$
\begin{aligned}
&l\cdot\|X^*(l,T,\omega)\|_2^2\\
&\leq \frac{1}{\lambda_{\min}}x^\top\Lambda x\exp\left(\theta(T-t)\right)\frac{\left(\frac{2\lambda_{\max}}{2l+\theta\lambda_{\max}}\right)^2}{\left(T-t+\frac{2\lambda_{\max}}{2l+\theta\lambda_{\max}}\right)^2}\prod_{t\leq\tau_i\leq s}(1+n^2)\frac{q(l,\tau_i)}{p(l,\tau_i)}. \quad (2.83)
\end{aligned}
$$

Furthermore,

$$\mathbb{E}_{t,x}\Big[\prod_{t\leq\tau_i\leq T}\frac{q(l,\tau_i)}{p(l,\tau_i)}\Big]<K_4<\infty$$

for some constant K_4 independent of l; thus for almost all $\omega\in\Omega$, there exists a constant $K(\omega)$ such that

$$\prod_{t\leq\tau_i(\omega)\leq T}\frac{q(l,\tau_i(\omega))}{p(l,\tau_i(\omega))}<K(\omega).$$

Therefore, Inequality (2.83) implies

$$\mathbb{E}_{t,x}\left[l\cdot\|X^*(l,T,\omega)\|_2^2\right]\xrightarrow{l\to\infty}0$$

and

$$l\cdot\|X^*(l,T,\omega)\|_2^2\xrightarrow{l\to\infty}0\quad\text{a.s.}$$

Finally, Lemma 2.4.7 implies that

$$\|\lim_{s\to T-}X^*(s)\|_2=0\quad\text{a.s.},$$

finishing the proof.

□

We can directly deduce compact convergence of the optimal trading intensity in the primary venue.

Corollary 2.4.8. *Let $t \in [0,T)$ and $x \in \mathbb{R}^n$ be the portfolio position an time t. Then*

$$\xi^*(l,\cdot,X^*(l,\cdot)) \longrightarrow \xi^*(\cdot,X^*(\cdot)) \quad \text{a.s. compactly on } [t,T)$$

as $l \to \infty$.

Proof. The assertion follows directly from the compact convergence results in Proposition 2.4.3 and Proposition 2.4.6. $\square$

We are now able to prove that u^* is indeed an admissible liquidation strategy. The main step towards this goal is accomplished by Proposition 2.4.6 (ii). It remains thus to show that u^* fulfills the moment conditions in Definition 2.1.1 (ii).

Theorem 2.4.9. *Let $t \in [0,T)$, $x \in \mathbb{R}^n$ and $u^* = (\xi^*, \eta^*)$ for ξ^* and η^* as in Equations* (2.74) *and* (2.75)*, respectively. Then u^* is an admissible liquidation strategy.*

Proof. Definition 2.1.1 (i) and (iii) are clear, (iv) follows form Proposition 2.4.6.

Let now $k \geq 1$ and $s \in [t,T)$. Similarly as in the proof of Proposition 2.4.6,

$$\mathbb{E}\Big[\Big(\prod_{t\leq\tau_i\leq s}(1+n^2)\frac{q(\tau_i)}{p(\tau_i)}\Big)^k\Big] \leq \mathbb{E}\Big[\Big(\prod_{t\leq\tau_i\leq T}(1+n^2)\frac{q(\tau_i)}{p(\tau_i)}\Big)^k\Big] < \infty. \tag{2.84}$$

Let now

$$\begin{aligned}
\tilde{p}(s) &:= \lim_{l\to\infty}\tilde{p}(l,s) = \frac{1}{T-s} - \frac{\theta}{2},\\
\tilde{q}(s) &:= \lim_{l\to\infty}\tilde{q}(l,s) = \frac{1}{T-s} + \sqrt{\alpha d_{\max}}.
\end{aligned}$$

Then

$$\begin{aligned}
r(s) &:= \tilde{q}(s)^k \exp\Big(-k\int_t^s \tilde{p}(v)dv\Big)\\
&= \Big(\frac{1}{T-s} + \sqrt{\alpha d_{\max}}\Big)^k \frac{(T-s)^k}{(T-t)^k}\exp\Big(\frac{k}{2}\theta(s-t)\Big)\\
&= (1-(T-s)\sqrt{\alpha d_{\max}})^k \frac{1}{(T-t)^k}\exp\Big(\frac{k}{2}\theta(s-t)\Big)
\end{aligned}$$

and thus

$$\int_t^T r(s) < \infty. \tag{2.85}$$

By Lemma 2.4.7, there exists a constant K such that

$$
\begin{aligned}
\mathbb{E}_{t,x}\Big[\int_t^T \|\xi(s,X^*(s))\|_2^4 ds\Big] &\leq \mathbb{E}_{t,x}\Big[\int_t^T q(s)^4 \|X^*(s)\|_2^4 ds\Big] \\
&\leq K\mathbb{E}_{t,x}\Big[\int_t^T \tilde{q}(s)^4 \exp\Big(-4\int_t^s \tilde{p}(v)dv\Big)\Big(\prod_{t\leq\tau_i\leq s}(1+n^2)\frac{q(\tau_i)}{p(\tau_i)}\Big)^4 ds\Big] \\
&\leq K\int_t^T \tilde{q}(s)^4 \exp\Big(-4\int_t^s \tilde{p}(v)dv\Big) ds\mathbb{E}_{t,x}\Big[(T-t)\Big(\prod_{t\leq\tau_i\leq T}(1+n^2)\frac{q(\tau_i)}{p(\tau_i)}\Big)^4\Big] \\
&\overset{(2.84),(2.85)}{<} \infty.
\end{aligned}
$$

By Corollary 2.2.6,

$$
\bar{C}(s) \leq nC(s)^{-1}
$$

and thus, as $\frac{q(s)}{p(s)}$ admits a continuous extension on $[t,T]$, there exists a constant $\bar{K}$ independent of s such that

$$
\|\bar{C}(s)C(s)\|_{2,2}^8 \leq n^8 \|C^{-1}\|_{2,2}^8 \|C\|_{2,2}^8 \leq \frac{n^8\lambda_{\max}^8 q(s)^8}{\lambda_{\min}^8 p(s)^8} \leq \bar{K}.
$$

Using Lemma 2.4.7 again, there exists a constant $\tilde{K}$ such that

$$
\begin{aligned}
\mathbb{E}_{t,x}\Big[\int_t^T \|\eta(s,X^*(s-)\|_2^8 ds\Big] &\leq \mathbb{E}_{t,x}\Big[\int_t^T \|\bar{C}(s)C(s)\|_2^8 \|X^*(s-)\|_2^8 ds\Big] \\
&\leq \bar{K}\mathbb{E}_{t,x}\Big[\int_t^T \exp\Big(-8\int_t^s p(v)dv\Big)\Big(\prod_{t\leq\tau_i\leq s}(1+n^2)\frac{q(\tau_i)}{p(\tau_i)}\Big)^8 ds\Big] \\
&\leq \bar{K}\int_t^T \exp\Big(-8\int_t^s p(v)dv\Big)\mathbb{E}_{t,x}\Big[(T-t)\Big(\prod_{t\leq\tau_i\leq s}(1+n^2)\frac{q(\tau_i)}{p(\tau_i)}\Big)^8\Big] \\
&\overset{(2.84)}{<} \infty.
\end{aligned}
$$

$\square$

2.4.3. Solution of the optimization problem

We are now ready to solve the Optimization Problem (OPT). By the results from the preceding sections, the limit of the solution of the Optimization Problem ($\widetilde{\text{OPT}}$) exists. In the following we verify that this limit is indeed the solution of the original Optimization Problem (OPT).

Theorem 2.4.10. *The value function of the Optimization Problem* (OPT) *is given by*

$$v(t,x) = x^\top C(t)x$$

for all $t \in [0,T)$, $x \in \mathbb{R}^n$ *and*

$$\lim_{s\to T-} v(s,x) = \begin{cases} 0 & \text{if } x = 0 \\ \infty & \text{else.} \end{cases}$$

The $\mathbb{P}\otimes\lambda$ *- almost surely unique optimal strategy is given by* $u^* = (\xi^*, \eta^*)$,

$$\begin{aligned} \xi^* &= \lim_{l\to\infty} \xi^*(l,t,x) = \Lambda^{-1}C(t)x, \\ \eta^* &= \lim_{l\to\infty} \eta^*(l,t,x) = \tilde{I}\bar{C}(t)C(t)x. \end{aligned}$$

Proof. We fix $t \in [0,T)$ and $x \in \mathbb{R}^n$. Note first that we have

$$v(t,x) \geq \lim_{l\to\infty} \tilde{v}(l,t,x). \tag{2.86}$$

For the converse inequality, let

$$A := \{\pi(T) = \pi(t)\}$$

be the set of scenarios without any dark pool execution in $[0,T]$ and $K : (l_0,\infty)\times\Omega \longrightarrow [0,\infty]$ be the following cost function:

$$\begin{aligned} K(l,\omega) := \int_t^T &\Big(\xi^*(l,s,X^*(l,s,\omega))^\top \Lambda \xi^*(l,s,X^*(l,s,\omega)) + \alpha X^*(l,s,\omega)^\top \Sigma X^*(l,s,\omega)\Big)\, ds \\ &+ l \cdot \|X^*(l,T,\omega)\|_2^2. \end{aligned}$$

Then $\mathbb{P}[A] > 0$ and for $\omega \in A$, $K(l,A) := K(l,\omega)$ is independent of the specific scenario ω almost surely. By optimality of $u^*(l)$ (Theorem 2.3.1), $K(l,A)$ is an upper bound for $K(l,\cdot)$ almost surely. As $\lim_{l\to\infty} \tilde{v}(l,t,x)$ is bounded and $\mathbb{P}[A] > 0$, there exists a constant K such that for all $l > l_0$

$$K(l,A) \leq K.$$

By the dominated convergence theorem this implies

$$\begin{aligned} \lim_{l\to\infty} \tilde{v}(l,t,x) &\overset{\text{Theorem 2.3.1}}{=} \lim_{l\to\infty} \mathbb{E}_{t,x}[K(l)] \\ &= \lim_{l\to\infty} \mathbb{E}_{t,x}\Big[\int_t^T \Big(\xi^*(l,s,X^*(l,s))^\top \Lambda \xi^*(l,s,X^*(l,s)) \\ &\qquad\qquad + \alpha X^*(l,s)^\top \Sigma X^*(l,s)\Big) ds + l \cdot \|X^*(l,T)\|_2^2\Big] \\ &= \mathbb{E}_{t,x}[\lim_{l\to\infty} K(l)] \end{aligned}$$

$$= \mathbb{E}_{t,x}\Big[\lim_{l\to\infty}\Big(\int_t^T \Big(\xi^*(l,s,X^*(l,s))^\top \Lambda \xi^*(l,s,X^*(l,s)) + \alpha X^*(l,s)^\top \Sigma X^*(l,s)\Big)ds + l\cdot \|X^*(l,T)\|_2^2\Big)\Big].$$

By Proposition 2.4.6 and Corollary 2.4.8, the limits exist, so Fatou's lemma yields

$$\begin{aligned}
\lim_{l\to\infty} \tilde{v}(l,t,x) &\geq \mathbb{E}_{t,x}\Big[\int_t^T \lim_{l\to\infty}\Big(\xi^*(l,s,X^*(l,s))^\top \Lambda \xi^*(l,s,X^*(l,s)) \\
&\qquad + \alpha X^*(l,s)^\top \Sigma X^*(l,s)\Big)ds + \lim_{l\to\infty} l\cdot \|X^*(l,T)\|_2^2\Big] \\
&= \mathbb{E}_{t,x}\Big[\int_t^T \Big(\xi^*(s,X^*(s))^\top \Lambda \xi^*(s,X^*(s)) + \alpha X^*(s)^\top \Sigma X^*(s)\Big)\,ds\Big] \\
&\geq v(t,x). \qquad (2.87)
\end{aligned}$$

The Inequalities (2.86) and (2.87) establish that u^* solves the Optimization Problem (OPT) and that the value function is given by v. For uniqueness, let $u=(\xi,\eta), \tilde{u} = (\tilde{\xi},\tilde{\eta}) \in \mathbb{A}(t,x)$ and $\mu\in(0,1)$. We define the convex combination $\bar{u} = (\bar{\xi},\bar{\eta})$:

$$\begin{aligned}
\bar{\xi}(s) &= \mu\xi(s) + (1-\mu)\tilde{\xi}(s), \\
\bar{\eta}(s) &= \mu\eta(s) + (1-\mu)\tilde{\eta}(s)
\end{aligned}$$

for $s\in[t,T)$. Thus,

$$X^{\bar{u}}(s) = \mu X^u(s) + (1-\mu)X^{\tilde{u}}(s)$$

and $\bar{u}\in\mathbb{A}(t,x)$. Notice that

$$\mathbb{P}\otimes\lambda\big[u\neq\tilde{u}\big] > 0 \quad \text{implies} \quad \mathbb{P}\otimes\lambda\big[\xi\neq\tilde{\xi}\big] > 0 \qquad (2.88)$$

as else

$$\mathbb{P}\big[\lim_{s\to T-} X^u(s) \neq \lim_{s\to T-} X^{\tilde{u}}(s)\big] > 0,$$

a contradiction to Definition 2.1.1 (iv). Hence,

$$\begin{aligned}
J(t,x,\bar{u}) &= \mathbb{E}_{t,x}\Big[\int_t^T f(\bar{\xi}(r), X^{\bar{u}}(r))dr\Big] \\
&\leq \mathbb{E}_{t,x}\Big[\int_t^T \mu f(\xi(r),X^u(r)) + (1-\mu) f(\tilde{\xi}(r),X^{\tilde{u}}(r))dr\Big] \qquad (2.89) \\
&= \mu J(t,x,u) + (1-\mu)J(t,x,\tilde{u}),
\end{aligned}$$

where Inequality (2.89) follows from the convexity of f. We have equality in Inequality (2.89) if and only if $u=\tilde{u}$ $\mathbb{P}\otimes\lambda$ - a.s. by strict convexity of f in the first argument and (2.88). □

2.5. Properties of the value function and the optimal strategy

Due to the analogy between the cost functionals in the discrete and the continuous-time setting, the value functions and the optimal strategies have corresponding properties.

In Section 2.5.1 we consider the single asset setting. Using Section 2.2.2, we derive closed form solutions for the value function and the optimal strategy. We prove similar monotonicity properties as for the corresponding objects in Section 1.3. The most interesting result is that the risk costs of the optimal strategy are decreasing in θ on the whole interval $(0,\infty)$ (Proposition 2.5.1 (v)), while in the discrete-time setting the risk costs are increasing for sufficiently small probability of execution p (cf. Proposition 1.3.9 (ii)). This effect disappears if we let the number of trading times $N+1$ tend to infinity. In Section 2.5.2 we discuss the portfolio case and analyze the bounds P and Q for the value function matrix C obtained in Section 2.4.1. By construction, P and Q are not sensitive to the signs of the positions in the assets. In the discrete-time setting we saw that different signs can yield different costs because of the correlation of the assets (cf. Example 1.3.10). Therefore, this property is rather undesirable, and we construct bounds that perform better and are sensitive to the signs of the positions. We illustrate the improvement by a numerical example.

2.5.1. Single asset liquidation

We let $n=1$ and set $\theta=\theta_1$. By Subsection 2.2.2, the solution of the Initial Value Problem (2.18) is given by

$$C(l,t)=\frac{\Lambda\tilde{\theta}}{2}\coth\Big(\frac{\tilde{\theta}}{2}(T-t)+\kappa(l)\Big)-\frac{\Lambda\theta}{2},$$

where

$$\kappa(l):=\operatorname{arcoth}\Big(\frac{\frac{2l}{\Lambda}+\theta}{\tilde{\theta}}\Big),$$
$$\tilde{\theta}:=\sqrt{\theta^2+\frac{4\alpha\Sigma}{\Lambda}}$$

for $\theta>0$ or $\alpha\Sigma>0$ and

$$C(l,t)=\frac{\Lambda}{T-t+\frac{\Lambda}{l}}$$

for $\theta=\alpha\Sigma=0$. In order to stress the dependence of the value function on the parameter θ, we define

$$\begin{aligned}C(t;\theta):=C(t)&=\lim_{l\to\infty}C(l,t)\\&=\begin{cases}\frac{\Lambda\tilde{\theta}}{2}\coth\Big(\frac{\tilde{\theta}}{2}(T-t)\Big)-\frac{\Lambda\theta}{2} & \text{if } \theta>0 \text{ or } \alpha\Sigma>0\\ \frac{\Lambda}{T-t} & \text{if } \theta=\alpha\Sigma=0,\end{cases}\end{aligned}\tag{2.90}$$

in particular

$$C(t;0) = \begin{cases} \sqrt{\alpha\Sigma\Lambda}\coth\left(\sqrt{\frac{\alpha\Sigma}{\Lambda}}(T-t)\right) & \text{if } \alpha\Sigma > 0 \\ \frac{\Lambda}{T-t} & \text{if } \alpha\Sigma = 0. \end{cases} \tag{2.91}$$

For simplicity, we consider the case $\alpha\Sigma > 0$ in the following. Thus, the value function v and the optimal strategy $u^* = (\xi^*, \eta^*)$ are given by

$$\begin{aligned} v(t,x) &= C(t;\theta)x^2 = \Big(\frac{\Lambda\tilde{\theta}}{2}\coth\Big(\frac{\tilde{\theta}}{2}(T-t)\Big) - \frac{\Lambda\theta}{2}\Big)x^2, \\ \xi^*(t,x;\theta) &= \frac{C(t;\theta)}{\Lambda}x = \Big(\frac{\tilde{\theta}}{2}\coth\Big(\frac{\tilde{\theta}}{2}(T-t)\Big) - \frac{\theta}{2}\Big)x, &(2.92) \\ \eta^*(t,x;\theta) &= \bar{C}(t;\theta)C(t;\theta)x = x. &(2.93) \end{aligned}$$

Similarly as in Section 1.3, it is always optimal to place the entire remainder of the position in the dark pool by Equation (2.93).

Let us now denote the optimal trading trajectory until execution in the dark pool by $\tilde{X}(\cdot;\theta)$, i.e., $\tilde{X}$ is the solution of the linear initial value problem

$$\begin{aligned} X' &= -\xi^*(\cdot, X;\theta) \\ X(0) &= x, \end{aligned}$$

where x is the initial asset position at time $t_0 = 0$. Then for $t \in [0,T)$,

$$\tilde{X}(t;\theta) = x\exp\Big(-\int_0^t \frac{C(s;\theta)}{\Lambda}ds\Big) = \frac{\sinh\left(\frac{\tilde{\theta}}{2}(T-t)\right)\exp\left(\frac{\theta}{2}t\right)}{\sinh\left(\frac{\tilde{\theta}}{2}T\right)}x. \tag{2.94}$$

In the following Proposition we show that many properties of the discrete-time case hold in the continuous case as well.

Proposition 2.5.1. *(i) For $t \in [0,T)$, $C(t;\theta)$ is strictly decreasing in θ.*

(ii) Let $t \in [0,T)$ and $x > 0$ be fixed. Then $\xi^*(t,x;\theta)$ is strictly decreasing in θ.

(iii) Assume that the initial asset position at time $t_0 = 0$ is given by $x > 0$. Then $\tilde{X}(t;\theta)$ is strictly increasing in θ for $t \in (0,T)$.

(iv) For an initial asset position $x > 0$ and $t \in (0,T)$, the expected asset position if the optimal strategy is applied,

$$\mathbb{E}[X^*(t;\theta)],$$

is strictly decreasing in θ.

(v) For $\alpha\Sigma > 0$, the risk costs

$$\alpha\Sigma \cdot \mathbb{E}\Big[\int_0^T X^*(t;\theta)^2 dt\Big]$$

are strictly decreasing in θ.

(vi) The impact costs

$$\Lambda \cdot \mathbb{E}\Big[\int_0^T \xi^*(t, X^*(t,\theta);\theta)^2 dt\Big]$$

are strictly decreasing in θ.

Proof. We let $t \in [0,T)$ and compute

$$\begin{aligned}
\frac{\partial}{\partial\theta} C(t;\theta) &= \frac{\Lambda\theta \coth\big(\frac{\tilde\theta}{2}(T-t)\big)}{2\tilde\theta} - \frac{\Lambda\theta(T-t)}{4\sinh^2\big(\frac{\tilde\theta}{2}(T-t)\big)} - \frac{\Lambda}{2} \\
&\leq \frac{\Lambda\theta\big(\cosh\big(\frac{\tilde\theta}{2}(T-t)\big)\sinh\big(\frac{\tilde\theta}{2}(T-t)\big) - \frac{\tilde\theta}{2}(T-t) - \sinh^2\big(\frac{\tilde\theta}{2}(T-t)\big)\big)}{2\tilde\theta\sinh^2\big(\frac{\tilde\theta}{2}(T-t)\big)} \\
&< 0
\end{aligned}$$

for $\theta > 0$ since

$$\begin{aligned}
&\cosh\big(\frac{\tilde\theta}{2}(T-t)\big)\sinh\big(\frac{\tilde\theta}{2}(T-t)\big) - \frac{\tilde\theta}{2}(T-t) - \sinh^2\big(\frac{\tilde\theta}{2}(T-t)\big) \\
&= \sinh\big(\frac{\tilde\theta}{2}(T-t)\big)\Big(\cosh\big(\frac{\tilde\theta}{2}(T-t)\big) - \sinh\big(\frac{\tilde\theta}{2}(T-t)\big)\Big) - \frac{\tilde\theta}{2}(T-t) \\
&= \Big(\frac{1-\exp(-\tilde\theta(T-t))}{2}\Big) - \frac{\tilde\theta}{2}(T-t) \\
&< 0
\end{aligned}$$

for $\theta > 0$ (note that $\frac{1}{2}(1-\exp(-2x)) - x < 0$ for $x > 0$). This establishes the first and the second assertion directly; the third assertion follows from the first equality in Equation (2.94).

For the proof of (iv), we note first that

$$\mathbb{E}[X^*(t;\theta)] = \mathbb{P}[\pi(t) = 0] \cdot \tilde X(t;\theta) = \frac{\sinh\big(\frac{\tilde\theta}{2}(T-t)\big)\exp\big(-\frac{\theta}{2}t\big)}{\sinh\big(\frac{\tilde\theta}{2}T\big)} x. \tag{2.95}$$

We compute for $\theta > 0$,

$$\begin{aligned}
\frac{\partial}{\partial\theta} X^*(t;\theta) = \frac{x}{\sinh^2\big(\frac{\tilde\theta}{2}(T-t)\big)} \Big(&\frac{\theta(T-t)}{2\tilde\theta}\sinh\big(\frac{\tilde\theta}{2}T\big)\cosh\big(\frac{\tilde\theta}{2}(T-t)\big)\exp\big(-\frac{\theta}{2}t\big) \\
&- \frac{t}{2}\sinh\big(\frac{\tilde\theta}{2}T\big)\sinh\big(\frac{\tilde\theta}{2}(T-t)\big)\exp\big(-\frac{\theta}{2}t\big) \\
&- \frac{\theta T}{2\tilde\theta}\cosh\big(\frac{\tilde\theta}{2}T\big)\sinh\big(\frac{\tilde\theta}{2}(T-t)\big)\exp\big(-\frac{\theta}{2}t\big)\Big)
\end{aligned}$$

$$
\begin{aligned}
&< \frac{\theta \exp\left(-\frac{\theta}{2}t\right)x}{2\tilde{\theta}\sinh^2\left(\frac{\tilde{\theta}}{2}(T-t)\right)}\Big((T-t)\sinh\left(\frac{\tilde{\theta}}{2}T\right)\cosh\left(\frac{\tilde{\theta}}{2}(T-t)\right) \\
&\qquad - T\cosh\left(\frac{\tilde{\theta}}{2}T\right)\sinh\left(\frac{\tilde{\theta}}{2}(T-t)\right)\Big) \\
&< 0
\end{aligned}
$$

by Lemma 1.3.8 (i) (cf. also Equation (1.44)), finishing the proof of (iv).

We have

$$
\mathbb{E}[X^*(t;\theta)^2] = \mathbb{P}[\pi(t)=0]\cdot\tilde{X}(t;\theta)^2 = \frac{\sinh\left(\frac{\tilde{\theta}}{2}(T-t)\right)}{\sinh\left(\frac{\tilde{\theta}}{2}T\right)}x^2.
$$

This term is differentiable and strictly decreasing in θ by Lemma 1.3.8 (note that $\tilde{\theta}$ is strictly increasing in θ). Thus, by Fubini's theorem,

$$
\frac{\partial}{\partial\theta}\mathbb{E}\Big[\int_0^T X^*(t;\theta)^2 dt\Big] = \frac{\partial}{\partial\theta}\int_0^T \mathbb{E}[X^*(t;\theta)^2]dt = \int_0^T \frac{\partial}{\partial\theta}\mathbb{E}[X^*(t;\theta)^2]dt < 0,
$$

establishing (v).

Finally, we note that

$$
\mathbb{E}[\xi^*(t,X^*(t;\theta);\theta)^2] = \mathbb{P}[\pi(t)=0]\cdot\frac{C(t;\theta)^2}{\Lambda^2}\tilde{X}(t;\theta)^2 = \frac{C(t;\theta)^2}{\Lambda^2}\mathbb{E}[X^*(t;\theta)^2]
$$

by Equation (2.95). This term is differentiable and strictly decreasing in θ as both terms are positive and strictly increasing in θ. Similarly as before, we deduce (vi). $\square$

Note that in the discrete-time setting, the risk costs are in general not decreasing in the probability of execution p (cf. Proposition 1.3.9 (ii)). By letting the number of trading times $N+1$ tend to infinity (and the length of each trading period tend to zero), appropriate scaling of the parameters Λ, Σ and p ensures that in the limit case the risk costs are strictly decreasing in p (cf. Section 2.6, in particular Remark 2.6.4).

2.5.2. Portfolio liquidation and improved bounds for the value function

As in the discrete-time setting of Section 1.3, it is in general not optimal for a trader liquidating more than one asset, to place the whole portfolio in the dark pool until execution. The optimal orders in the dark pool depend strongly on the correlation of the assets. The properties we identified in the discrete-time setting transfer to corresponding properties in the continuous-time setting. As we discussed this issue in detail in Section 1.3.4 for the discrete-time case, we will not go into too much detail here. We want to remark, however, that the discrete-time setting is somewhat more general as it allows for a dependencies between the dark pool liquidities of the assets (cf. Section 1.3.4). Throughout Chapter 2 we assumed that the Poisson processes which govern dark pool execution are independent. Therefore, the discussion about dependencies of the dark pool liquidities in Example 1.3.11 does not apply here.

As we are not able to obtain closed form solutions for the value function and the optimal strategy for $n \geq 2$, we analyze the bounds for the value function $C(l, t)$ of the Optimization Problem ($\widetilde{\text{OPT}}$) from Theorem 2.2.8:

$$0 < P(l, t) \leq \sqrt{\Lambda^{-1}} C(l, t) \sqrt{\Lambda^{-1}} \leq Q(l, t)$$

or equivalently

$$0 < \sqrt{\Lambda} P(l, t) \sqrt{\Lambda} \leq C(l, t) \leq \sqrt{\Lambda} Q(l, t) \sqrt{\Lambda}.$$

These bounds transfer to bounds for the value function matrix $C(t) = \lim_{l \to \infty} C(l, t)$ of the Optimization Problem (OPT) (cf. Theorem 2.4.4):

$$0 < \lim_{l \to \infty} P(l, t) = P(t) \leq \sqrt{\Lambda^{-1}} C(t) \sqrt{\Lambda^{-1}} \leq \lim_{l \to \infty} Q(l, t) = Q(t)$$

or equivalently

$$0 < \sqrt{\Lambda} P(t) \sqrt{\Lambda} \leq C(t) \leq \sqrt{\Lambda} Q(t) \sqrt{\Lambda}.$$

We require these bounds in several proofs (e.g., for establishing bounds of the norms of the processes $\|X^*(l, t)\|_2$ and $\|X^*(t)\|_2$, respectively, which are important for the proof of the main result, Theorem 2.4.10). It proved useful to construct these bounds in such a way that they were multiples of the identity matrix:

$$P(l, t) = p(l, t) I, \quad Q(l, t) = q(l, t) I, \quad P(t) = p(t) I, \quad Q(t) = q(t) I \tag{2.96}$$

for explicitly known functions $p(l), q(l) : [0, T] \to \mathbb{R}$, $p, q : [0, T) \to \mathbb{R}$, so that they commute with every matrix $A \in \mathbb{R}^{n \times n}$. Thus, the lower respectively upper bound for the value function $x^\top \Lambda P x$ respectively $x^\top \Lambda Q x$ are *independent* of the signs of x_i, $i = 1, \ldots, n$, and thus ignore the diversification of the portfolio. In the following we compute upper and lower bounds for $C(t)$ in closed form that improve the bounds $P(t)$ and $Q(t)$ by giving up Property (2.96). These bounds are sensitive for the signs of the asset positions x_i. To this end, we replace the Initial Value Problem (2.18) by

$$\begin{aligned} C' &= C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha \Sigma \\ C(T) &= l \Lambda. \end{aligned} \tag{2.97}$$

Let us assume that $\alpha \Sigma > 0$ for simplicity. Recall that $D = \sqrt{\Lambda^{-1}} \Sigma \sqrt{\Lambda^{-1}}$ and consider the initial value problems

$$\begin{aligned} P' &= P^2 + \theta P - \alpha D \\ P(T) &= l I \end{aligned} \tag{2.98}$$

and

$$\begin{aligned} Q' &= P^2 - \alpha D \\ Q(T) &= l I. \end{aligned} \tag{2.99}$$

By a slight abuse of notation, we denote the solution of (2.97) by $C(l,t)$. By Theorem 2.2.2, the Initial Value Problems (2.98) and (2.99) possess solutions $\bar{P}(l,t)$ respectively $\bar{Q}(l,t)$ on the interval $(-\infty,T]$ for large enough l. Similarly as in the proof of Theorem 2.2.8, we obtain that the Initial Value Problem (2.97) possesses a solution on $(-\infty,T]$ for large enough l and that

$$0 < P(l,t) \leq \bar{P}(l,t) \leq \sqrt{\Lambda^{-1}}C(l,t)\sqrt{\Lambda^{-1}} \leq \bar{Q}(l,t) \leq Q(l,t).$$

We obtain monotonicity of $(C(l,t))_l$ by the same proof as in Lemma 2.4.2 (where we proved monotonicity for initial values lI). Therefore, $(C(l,t))_l$ converges, and the limit must be the same as before:

$$C(t) = \lim_{l\to\infty} C(l,t).$$

Thus, provided that

$$\bar{P}(t) := \lim_{l\to\infty} \bar{P}(l,t) \quad \text{and} \quad \bar{Q}(t) := \lim_{l\to\infty} \bar{Q}(l,t)$$

exist,

$$0 < P(t) \leq \bar{P}(t) \leq \sqrt{\Lambda^{-1}}C(t)\sqrt{\Lambda^{-1}} \leq \bar{Q}(t) \leq Q(t)$$

or equivalently

$$0 < \sqrt{\Lambda}P(t)\sqrt{\Lambda} \leq \sqrt{\Lambda}\bar{P}(t)\sqrt{\Lambda} \leq C(t) \leq \sqrt{\Lambda}\bar{Q}(t)\sqrt{\Lambda} \leq \sqrt{\Lambda}Q(t)\sqrt{\Lambda}. \tag{2.100}$$

In the following we solve the Initial Value Problems (2.98) and (2.99) explicitly using matrix functions (see, e.g., the book by Horn and Johnson [1991]) and compute the limits $\bar{P}(t)$ and $\bar{Q}(t)$. By the above discussion, we obtain that Equation (2.100) holds.

Proposition 2.5.2. *Let*

$$l_1 := \max_{i=1\dots,n} \{\sqrt{\alpha d_i}\},$$

where $d_1,\dots,d_n > 0$ are the eigenvalues of D. For $l > l_1$, the solutions of the Initial Value Problems (2.98) *and* (2.99) *are given by*

$$\bar{P}(l,t) = \sqrt{\bar{D}}\coth\Big(\sqrt{\bar{D}}(T-t) + \operatorname{arcoth}\big((l+\frac{\theta}{2})\sqrt{\bar{D}}^{-1}\big)\Big) - \frac{\theta}{2}I, \tag{2.101}$$

$$\bar{Q}(l,t) = \sqrt{\tilde{D}}\coth\Big(\tilde{D}(T-t) + \operatorname{arcoth}\big(l\sqrt{\tilde{D}}^{-1}\big)\Big), \tag{2.102}$$

respectively, where

$$\bar{D} := \frac{\theta^2}{4}I + \alpha D, \quad \tilde{D} := \alpha D.$$

In particular,

$$\bar{P}(t) = \lim_{l\to\infty} \bar{P}(l,t) = \bar{D}\coth\big(\bar{D}(T-t)\big)\Big) - \frac{\theta}{2}I, \tag{2.103}$$

$$\bar{Q}(t) = \lim_{l\to\infty} \bar{Q}(l,t) = \tilde{D}\coth\big(\tilde{D}(T-t)\big). \tag{2.104}$$

Proof. We prove Equation (2.101). Equation (2.102) follows directly by setting $\theta = 0$. Equations (2.103) and (2.104) are a direct consequence.

Let us consider the scalar initial value problems ($i = 1, \dots, n$)

$$\begin{aligned} p' &= p^2 + \theta p - \alpha d_i \\ p(T) &= l. \end{aligned}$$

By Section 2.2.2, the solution $p_i(l,\cdot)$ of the i^{th} initial value problem exists on $(-\infty, T]$ and is given by

$$p_i(l,t) = \sqrt{\frac{\theta^2}{4} + \alpha d_i} \coth\Big(\sqrt{\frac{\theta^2}{4} + \alpha d_i}(T-t) + \operatorname{arcoth}\Big(\frac{l + \frac{\theta}{2}}{\sqrt{\frac{\theta^2}{4} + \alpha d_i}}\Big)\Big) - \frac{\theta}{2}.$$

Recall that D is symmetric. By the spectral theorem, there exists an orthogonal matrix $U \in \mathbb{R}^{n \times n}$ with

$$U^\top D U = \operatorname{diag}(d_1, \dots, d_n).$$

We obtain that

$$\begin{aligned} \frac{\partial}{\partial t} U \operatorname{diag}(p_i(l,t)) U^\top &= U \operatorname{diag}\big(\frac{\partial}{\partial t} p_i(l,t)\big) U^\top \\ &= U \operatorname{diag}(p_i(l,t)^2) U^\top + \theta U \operatorname{diag}(p_i(l,t)) U^\top - \alpha U \operatorname{diag}(d_i) U^\top \\ &= \big(U \operatorname{diag}(p_i(l,t)) U^\top\big)^2 + \theta\big(U \operatorname{diag}(p_i(l,t)) U^\top\big) - \alpha D \end{aligned}$$

and therefore

$$\begin{aligned} \bar{P}(l,t) &= U \operatorname{diag}(p_i(l,t)) U^\top \\ &= U \operatorname{diag}\big(\sqrt{\theta^2/4 + \alpha d_i}\big) U^\top U \operatorname{diag}\big(\coth(\dots)\big) U^\top - \frac{\theta}{2} I \\ &= \sqrt{\bar{D}} \coth\Big(U \operatorname{diag}\big(\sqrt{\theta^2/4 + \alpha d_i}\big) U^\top (T-t) \\ &\qquad + \operatorname{arcoth}\big(U \operatorname{diag}\big(\frac{l + \theta/2}{\sqrt{\theta^2/4 + \alpha d_i}}\big) U^\top\big)\Big) - \frac{\theta}{2} I \\ &= \sqrt{\bar{D}} \coth\Big(\sqrt{\bar{D}}(T-t) + \operatorname{arcoth}\big((l + \frac{\theta}{2})\sqrt{\bar{D}}^{-1}\big)\Big) - \frac{\theta}{2} I. \end{aligned}$$

□

As the bounds $x^\top \sqrt{\Lambda} \bar{P} \sqrt{\Lambda} x$ and $x^\top \sqrt{\Lambda} \bar{Q} \sqrt{\Lambda} x$ are sensitive for the sign of the asset position, we expect them to be a significant improvement over $x^\top \Lambda P x$ respectively $x^\top \Lambda Q x$. We analyze the improvement in the following numerical example.

Example 2.5.3. *We consider the continuous-time analog of the discrete-time setting of*

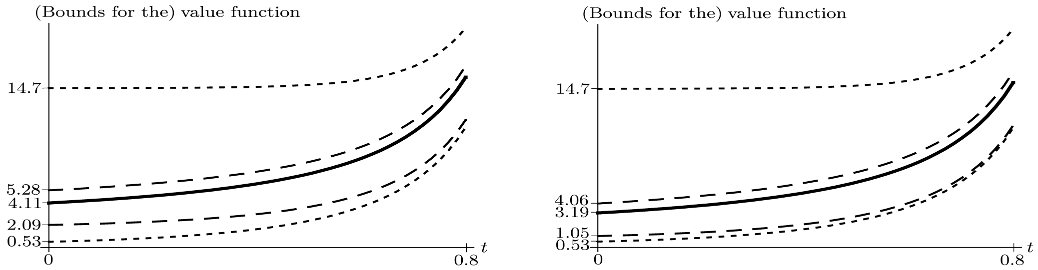


Figure 2.1.: Value function $v(t,x) = x^\top C(t)x$ and bounds for the value function for $t \in [0, 0.8]$, $x = (1,1)^\top$ (left picture), $x = (1,-1)^\top$ (right picture). In both pictures the solid line represents the value function, dashed lines represent the improved upper respectively lower bounds and dotted lines represent the bounds form Section 2.4.1.

Example 1.3.10. Let

$$\Lambda = \begin{pmatrix} 3 & 0 \\ 0 & 0.2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 1 & 0.9 \\ 0.9 & 1 \end{pmatrix}, \quad \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 3 \end{pmatrix}, \quad T = 1.$$

As $x^\top \Lambda P x$ and $x^\top \Lambda Q x$ are independent of the signs of x_i ($i = 1, 2$), these bounds are the same for the well-diversified portfolio $(1,1)^\top$ and the poorly diversified portfolio $(1,-1)^\top$ (note the high correlation of the two stocks, cf. also Section 1.3.4) although the value function differs significantly as we demonstrate below.

We solve the Initial Value Problem (2.18) *numerically and compute the value function $v(t,\cdot)$ for $x = (1,1)^\top$ and $y = (1,-1)^\top$, respectively. We obtain*

$$3.19 = v(0,y) < v(0,x) = 4.11, \quad 15.29 = v(0.8,y) < v(0.8,x) = 15.70.$$

The value function at x and y differs significantly at both points in time $t = 0$ and $t = 0.8$. Furthermore, $\bar{P}$ and $\bar{Q}$ perform significantly better than P and Q, respectively. The relative differences between $v(t,x)$ and $v(t,y)$ as well as between P, Q and $\bar{P}$, $\bar{Q}$ are much smaller for $t = 0.8$ than for $t = 0$. The reason is that for very short time horizons, risk costs are not as significant and impact costs outweigh risk costs. Thus, the diversification of the portfolio and the fact that P and Q ignore the sign of the asset positions are less important for $t = 0.8$.

We illustrate this in Figure 2.1. In both pictures the thick solid line denotes the value function. Dashed lines correspond to the improved bounds and dotted lines correspond to the original bounds. As expected, the bounds $x^\top \sqrt{\Lambda}\bar{P}\sqrt{\Lambda}x$ and $x^\top \sqrt{\Lambda}\bar{Q}\sqrt{\Lambda}x$ perform significantly better then $x^\top \Lambda P x$ respectively $x^\top \Lambda Q x$ in both cases

2.6. Discretization

In Chapter 1 we introduced a discrete-time market model for liquidating a large portfolio both by trading in the primary exchange and by trading in a dark pool. For a specification of this model where the price impact in the market is linear and temporary, we obtained a recursive scheme for solving the associated optimization problem in Section 1.3. We are thus able to compute the value function and the optimal strategy. The construction of the corresponding continuous-time market model in Chapter 2 suggests that there should be a close connection between the value function of the discrete-time model and the value function of the continuous-time model. We believe that the discrete-time value function converges to the continuous-time value function as the number of trading times N tends to infinity.

In order to obtain such a convergence result, we have to scale the parameters of the discrete-time setting appropriately. For example, for larger N, i.e., shorter trading periods, the price impact per trading period should be larger, while the covariance matrix and the probability of execution should be smaller.

In Section 2.6.1 we specify the scaling of the parameters and the form of the convergence we expect. In Section 2.6.2 we prove convergence for the single asset setting. For this case, we have closed form solutions for the value function both for the discrete-time model and for the continuous-time model. In Section 2.6.3 we discuss the case of liquidating a portfolio and give heuristic reasons why we think that the convergence result also holds for general $n \in \mathbb{N}$. In Section 2.6.4 we show that convergence of the value function transfers to convergence of the optimal strategy.

2.6.1. Model description

In order to compare the discrete-time model and the continuous-time model, we first have to introduce notations that avoid ambiguities. For example, the impact matrix Λ in the discrete-time setting refers to a slightly different object than in the continuous-time setting.

We will keep the notations for the continuous-time setting of Chapter 2. $\Lambda > 0$ denotes the price impact matrix and $\Sigma \geq 0$ denote the covariance matrix for the unit interval $[0, 1]$. Furthermore, $\theta_1, \ldots, \theta_n$ denote the intensities of the n independent Poisson processes $\pi_1, \ldots, \pi_n$. We assume that

$$\theta_i > 0$$

for simplicity.

Let us now consider the discrete-time model of Section 1.3 with $N + 1$ trading times. We make the following two simplifying specifications.

Assumption 2.6.1. *(i) For $j = 0, \ldots, N$, the trading intervals $[t_j, t_{j+1})$ have equal length*

$$t_{j+1} - t_j = \frac{T}{N}.$$

(ii) The dark pool liquidities for the n assets are independent.

Assumption 2.6.1 (ii) is necessary as we only consider independent dark pools in the continuous-time setting. We also assume implicitly that the last trade takes place in the interval $[t_N, t_{N+1})$ rather than at time t_N and that the trading horizon is $[0, T + \frac{1}{N})$ with no dark pool orders allowed in $[T, T + \frac{1}{N})$.

We scale the parameters Λ, Σ, $\hat{P}$ (cf. Equation (1.22)) in the following way.

(i) Given the n-dimensional independent Poisson process π, we replace the matrix $\hat{P}$ by

$$\hat{P}(N) = \operatorname{diag}\left(\hat{p}_{i,i}(N)\right), \quad \hat{p}_{i,i}(N) := \frac{\theta_i T}{N}.$$

Recall that $\hat{p}_{i,i}(N)$ is the probability of execution of orders of the i^{th} asset in the dark pool in a given time interval $[t_j, t_{j+1})$. Consequently, the expected number of executions for the i^{th} asset in $[0, T]$ is given by

$$\theta_i T$$

(independent of $N \in \mathbb{N}$) both in the discrete-time and in the continuous-time setting.

(ii) We replace the price impact matrix by

$$\Lambda(N) := \frac{\Lambda(N+1)}{T}.$$

Liquidating a portfolio $X \in \mathbb{R}^n$ at trading times $0 = t_0 < \cdots < t_N = T$ in equally large blocks in the primary venue yields the price impact costs

$$\sum_{i=0}^{N} \frac{1}{(N+1)^2} X^\top \frac{\Lambda(N+1)}{T} X = \frac{1}{T} X^\top \Lambda X.$$

Liquidating the same portfolio continuously in $[0, T]$ at constant trading rate yields thus the same price impact costs:

$$\int_0^T \frac{1}{T^2} X^\top \Lambda X dt = \frac{1}{T} X^\top \Lambda X.$$

(iii) We replace the covariance matrix by

$$\Sigma(N) := \frac{\Sigma T}{N+1},$$

i.e., we assume implicitly that the variances and the covariances of the price processes $\tilde{P}(t_k)$ are linear in time. This is, e.g., the case if the price increments are identically distributed (recall that we assumed independence of the price increments in Chapter 1).

Also, for the value function matrix and the optimal strategy, we keep the notation for the continuous-time setting and denote it by $C(t)$ for $t \in [0, T)$. For the discrete-time setting with $N+1$ trading times, we denote the value function matrix C and the matrices $\check{C}$ and D (cf. Notation 1.3.3 and Equation (1.26)) at a given time $t = t_k$ by

$$C(t,N) = (c_{i,j}(t,N))_{i,j,=1,\ldots,n}, \quad \check{C}(t,N) = (\check{c}_{i,j}(t,N))_{i,j,=1,\ldots,n}$$

and

$$D(t,N) = (d_{i,j}(t,N))_{i,j,=1,\ldots,n},$$

respectively.

The goal is now to show the following conjecture.

Conjecture 2.6.2. *Let $t \in [0,T)$ and assume that there is a sequence $t(N) \in [0,T]$ such that*

$$t(N) \in \{0, \tfrac{T}{N}, \tfrac{2T}{N}, \ldots, \tfrac{NT}{N} = T\}$$

and

$$\lim_{N\to\infty} t(N) = t.$$

Then (element-wise)

$$\lim_{N\to\infty} C(t(N),N) = C(t). \tag{2.105}$$

2.6.2. Single asset liquidation

In Section 1.3.3 we solved the discrete-time Optimization Problem ($\mathrm{OPT_{dis}}$) in closed form for $n = 1$. Let $N \in \mathbb{N}$, $t(N) := t_i(N) := \frac{iT}{N}$ $(i = 0,\ldots,N)$ and $p(N) = \frac{\theta T}{N}$ be the probability of execution. Using the notations from Section 2.6.1, the value function matrix C is given by (cf. Proposition 1.3.5)

$$C(t(N),N) = \frac{\Lambda(N)}{1-p(N)}\Big(\frac{\sqrt{1-p(N)}\sinh\Big(\kappa(p(N))(N+2-i)\Big)}{\sinh\Big(\kappa(p(N))(N+1-i)\Big)} - 1\Big)$$

for

$$\kappa(p(N)) = \operatorname{arcosh}\Big(\frac{\sqrt{1-p(N)}}{2}\big(\frac{\alpha\Sigma(N)}{\Lambda(N)} + 1 + \frac{1}{1-p(N)}\big)\Big).$$

We also derived the solution of the continuous-time Optimization Problem (OPT) (cf. Equation (2.90)) in closed form. For $t \in [0,T)$, the value function matrix is given by

$$C(t) = \frac{\Lambda\tilde{\theta}}{2}\coth\Big(\frac{\tilde{\theta}}{2}(T-t)\Big) - \frac{\Lambda\theta}{2}$$

for

$$\tilde{\theta} = \sqrt{\theta^2 + \frac{4\alpha\Sigma}{\Lambda}}.$$

Using these closed form solutions, we can prove Conjecture 2.6.2 for the case $n = 1$.

Theorem 2.6.3. *For $n = 1$, Conjecture 2.6.2 holds.*

Proof. Let $t \in [0, T)$ and assume that there is a sequence $t(N) \in [0, T]$ such that for $i(N) \in \{0, \ldots, N\}$,

$$t(N) = \frac{i(N)T}{N} = t + \epsilon(N)$$

with

$$\lim_{N \to \infty} \epsilon(N) = 0.$$

Using the addition formula for sinh, we obtain

$$\begin{aligned} C(t(N), N) &= \frac{\Lambda(N)}{\sqrt{1-p(N)}} \Big(\frac{\sinh\big(\kappa(p(N))(N+1-i)\big) \cosh\big(\kappa(p(N))\big)}{\sinh\big(\kappa(p(N))(N+1-i)\big)} \\ &\quad + \frac{\cosh\big(\kappa(p(N))(N+1-i)\big) \sinh\big(\kappa(p(N))\big)}{\sinh\big(\kappa(p(N))(N+1-i)\big)} - \frac{1}{\sqrt{1-p(N)}} \Big) \\ &= \frac{\Lambda}{T\sqrt{1-p(N)}} \Big((N+1) \Big(\cosh\big(\kappa(p(N))\big) - \frac{1}{\sqrt{1-p(N)}} \Big) \\ &\quad + (N+1) \sinh\big(\kappa(p(N))\big) \coth\big(\kappa(p(N))(N+1-i)\big) \Big). \end{aligned} \tag{2.106}$$

We have

$$\begin{aligned} &(N+1)\Big(\cosh\big(\kappa(p(N))\big) - \frac{1}{\sqrt{1-p(N)}} \Big) \\ &\quad = (N+1) \Big(\frac{\sqrt{1-p(N)}}{2} \big(\frac{\alpha\Sigma(N)}{\Lambda(N)} + 1 + \frac{1}{1-p(N)} \big) - \frac{1}{\sqrt{1-p(N)}} \Big) \\ &\quad = \frac{\sqrt{1-\frac{\theta T}{N}}\alpha\Sigma T^2}{2(N+1)\Lambda} + \frac{N+1}{2} \Big(\sqrt{1-\frac{\theta T}{N}} - \frac{1}{\sqrt{1-\frac{\theta T}{N}}} \Big) \\ &\quad \longrightarrow -\frac{\theta T}{2} \end{aligned} \tag{2.107}$$

as $N \to \infty$. Moreover,

$$\begin{aligned} (N+1)^2 \sinh^2\big(\kappa(p(N))\big) &= (N+1)^2 \big(\cosh^2\big(\kappa(p(N))\big) - 1 \big) \\ &= \frac{1}{4}(N+1)^2 \Big(\Big(\sqrt{1-p(N)} + \frac{1}{\sqrt{1-p(N)}} \frac{\alpha\Sigma(N)\sqrt{1-p(N)}}{\Lambda(N)} \Big)^2 - 4 \Big) \\ &= \frac{1}{4}(N+1)^2 \Big((1 - \frac{\theta T}{N}) - 2 + \frac{1}{1-\frac{\theta T}{N}} + 2\frac{\alpha\Sigma T^2 \sqrt{1-\frac{\theta T}{N}}(2-\frac{\theta T}{N})}{(N+1)^2\Lambda} \\ &\quad + \frac{\alpha^2\Sigma^2T^4(1-\frac{\theta T}{N})}{(N+1)^4\Lambda^2} \Big) \\ &\longrightarrow \frac{\theta^2T^2}{4} + \frac{\alpha\Sigma T^2}{\Lambda} = \frac{T^2\tilde{\theta}^2}{4} \end{aligned}$$

as $N \to \infty$ and hence

$$(N+1)\sinh\big(\kappa(p(N))\big) \longrightarrow \frac{T\tilde{\theta}}{2} \tag{2.108}$$

and (note that $\lim_{N\to\infty} \kappa(p(N)) = 0$)

$$N\kappa(p(N)) = N\sinh\big(\kappa(p(N))\big)\frac{\kappa(p(N))}{\sinh\big(\kappa(p(N))\big)} \longrightarrow \frac{T\tilde{\theta}}{2} \tag{2.109}$$

as $N \to \infty$. We note that

$$N+1-i = \frac{N}{T}(T + \frac{T}{N} - \frac{iT}{N}) = \frac{N}{T}(T - t + \tilde{\epsilon}(N))$$

for $\lim_{N\to\infty} \tilde{\epsilon}(N) = 0$ and deduce by Equation (2.106) (using (2.107), (2.108) and (2.109)) that

$$\lim_{N\to\infty} C(t(N), N) = -\frac{\Lambda\theta}{2} + \frac{\Lambda\tilde{\theta}}{2}\coth\big(\frac{\tilde{\theta}}{2}(T-t)\big)$$

as required. □

Remark 2.6.4. *In Section 1.3.3 we showed that in the discrete-time setting, the risk costs are increasing in $p(N)$ for*

$$p(N) < \frac{\alpha\Sigma(N)}{\Lambda(N) + \alpha\Sigma(N)} \tag{2.110}$$

(cf. Proposition 1.3.9), while in the continuous-time setting, they are decreasing in θ for $\theta > 0$ (cf. Proposition 2.5.1). Considering Theorem 2.6.3, this is surprising at first sight. We want to remark here that in the limit case of the discrete-time setting, the risk costs are decreasing on the whole interval $(0, 1)$ as Inequality (2.110) is equivalent to

$$\theta < \frac{NT\alpha\Sigma}{(N+1)^2\Lambda + \alpha\Sigma T^2} \longrightarrow 0$$

as $N \to \infty$.

2.6.3. Portfolio liquidation

For $n \geq 2$, we neither have closed form solutions for the discrete-time optimization problem nor for the continuous-time optimization problem. Therefore, we cannot prove Conjecture 2.6.2 in a similar way as for $n = 1$.

For $N \in \mathbb{N}$, $t = \frac{iT}{N}$, $i \in \{0, \dots, N\}$, the discrete-time value function matrix is given recursively by $C(T, N) = \Lambda(N) + \alpha\Sigma(N)$,

$$C(t,N) = \alpha\Sigma(N) + D(t + \tfrac{T}{N}, N) - D(t + \tfrac{T}{N}, N)\Big(\Lambda(N) + D(t + \tfrac{T}{N}, N))\Big)^{-1} D(t + \tfrac{T}{N}, N),$$

where

$$D(t+\tfrac{T}{N},N) = C(t+\tfrac{T}{N},N) - C(t+\tfrac{T}{N},N)\hat{P}(N)\check{C}(t+\tfrac{T}{N},N)^{-1}\hat{P}(N)C(t+\tfrac{T}{N},N)$$

(cf. Equation (1.25)).

For the continuous-time setting, the value function matrix solves the matrix differential equation

$$C' = C^\top \Lambda^{-1} C + C^\top \tilde{C} C - \alpha\Sigma, \tag{2.111}$$

where

$$\tilde{C} := \operatorname{diag}\left(\tfrac{\theta_i}{c_{i,i}}\right)$$

with boundary condition

$$\lim_{t\to T-} c_{\min}(t) = \infty \tag{2.112}$$

(cf. Theorem 2.4.4).

It turns out that it is difficult to connect these two objects. One problem is that (2.111) is not a matrix Riccati equation, and therefore it is not clear that the value function matrix is the *unique* solution of (2.111) with boundary condition (2.112) (cf. Remark 2.4.5). In the following we give heuristic reasons why we believe that Conjecture 2.6.2 is true for general $n \in \mathbb{N}$. It is left for future research to give a rigorous proof.

We note first that by the independence of the dark pools, we obtain

$$\begin{aligned}\check{c}_{i,j}(t,N) &= \begin{cases} c_{i,i}(t,N)\hat{p}_{i,i}(N) & \text{if } i=j\\ c_{i,j}(t,N)\hat{p}_{i,i}(N)\hat{p}_{j,j}(N) & \text{if } i\neq j\end{cases}\\ &= \begin{cases} \frac{c_{i,i}(t,N)\theta_i T}{N} & \text{if } i=j\\ \frac{c_{i,j}(t,N)\theta_i\theta_j T^2}{N^2} & \text{if } i\neq j.\end{cases}\end{aligned} \tag{2.113}$$

We define the matrix $\hat{C}(t,N) = (\hat{c}_{i,j}(t,N))_{i,j=1,\dots,n}$ by

$$\hat{C}(t,N) := \hat{P}(N)^{-1}\check{C}(t,N)\hat{P}(N)^{-1}.$$

Note that $\hat{C}$ is well-defined since we assumed $\theta_i > 0$ and that

$$\begin{aligned}\hat{c}_{i,j}(t,N) &= \begin{cases} \frac{c_{i,i}(t,N)}{\hat{p}_{i,i}(N)} & \text{if } i=j\\ c_{i,j}(t,N) & \text{if } i\neq j\end{cases}\\ &= \begin{cases} \frac{c_{i,i}(t,N)N}{\theta_i T} & \text{if } i=j\\ c_{i,j}(t,N) & \text{if } i\neq j.\end{cases}\end{aligned} \tag{2.114}$$

We will make the following assumption.

Assumption 2.6.5. *Let $t \in [0,T)$ and consider a sequence $t(N) \in [0,T]$ such that $t(N) \in \{0, \frac{T}{N}, \dots, T\}$ and*

$$\lim_{N\to\infty} t(N) = t.$$

Then, there exists a positive definite matrix $F(t) = (f_{i,j}(t))_{i,j=1,\dots,n} \in \mathbb{R}^{n\times n}$ *such that*

$$\lim_{N\to\infty} C(t(N), N) = F(t). \tag{2.115}$$

The following lemma implies the existence of a convergent subsequence.

Lemma 2.6.6. *Let* $t \in [0, T)$ *and assume that there is a sequence* $t(N) \in [0, T]$ *such that* $t(N) \in \{0, \frac{T}{N}, \dots, T\}$ *and*

$$\lim_{N\to\infty} t(N) = t.$$

Then there exists a matrix $G \in \mathbb{R}^{n\times n}$ *such that*

$$C(t(N), N) \leq G \quad \textit{for all } N \in \mathbb{N}.$$

Proof. Let $x \in \mathbb{R}^n$ be the portfolio position at time $t(N)$. Assume that $t(N) = \frac{i(N)T}{N}$ and define the following admissible strategy:

$$x(t_j, N) := \frac{1}{N - i(N) + 1} x, \quad y(t_j, N) := 0, \quad j = i, \dots, N.$$

This strategy yields at least the costs of the optimal strategy. For $X(t_{i(N)}, N) = x$, $X(t_j, N) = X(t_{j-1}, N) - x(t_{j-1}, N)$, we obtain

$$\begin{aligned} x^\top C(t(N), N)x &\leq \sum_{j=i(N)}^{N} x(t_j, N)^\top \Lambda(N) x(t_j, N) + \alpha \sum_{j=i(N)}^{N} X(t_j, N)^\top \Sigma(N) X(t_j, N) \\ &\leq \frac{N+1}{N - i(N) + 1} x^\top \Lambda x + \alpha x^\top \Sigma x \\ &\longrightarrow \frac{T}{T-t} x^\top \Lambda x + \alpha x^\top \Sigma x \end{aligned}$$

as $N \longrightarrow \infty$. □

We obtain the following result.

Proposition 2.6.7. *Let* $t \in [0, T)$ *and* $t(N) \in [0, T]$ *such that* $t(N) \in \{0, \frac{T}{N}, \dots, T\}$ *and*

$$\lim_{N\to\infty} t(N) = t.$$

If Assumption 2.6.5 holds, then

$$\lim_{N\to\infty} \frac{C\big(t(N) + \frac{T}{N}, N\big) - C\big(t(N), N\big)}{T/N} = F(t)\Lambda^{-1}F(t) + F(t)\tilde{F}(t)F(t) - \alpha\Sigma.$$

Proof. Note first that

$$C\big(t(N) + \tfrac{T}{N}, N\big) - C\big(t(N), N\big)$$

$$
\begin{aligned}
&= -\alpha\Sigma(N) + C\big(t(N)+\tfrac{T}{N},N\big)\hat{C}\big(t(N)+\tfrac{T}{N},N\big)^{-1}C\big(t(N)+\tfrac{T}{N},N\big)\\
&\quad - D\big(t(N)+\tfrac{T}{N},N\big)\Big(\Lambda(N)+D\big(t(N)+\tfrac{T}{N},N\big)\Big)^{-1}D\big(t(N)+\tfrac{T}{N},N\big)
\end{aligned}
$$

(recall that

$$D\big(t(N)+\tfrac{T}{N},N\big) = C\big(t(N)+\tfrac{T}{N},N\big) - C\big(t(N)+\tfrac{T}{N},N\big)\hat{C}\big(t(N)+\tfrac{T}{N},N\big)^{-1}C\big(t(N)+\tfrac{T}{N},N\big)$$

and note that $\hat{C}^{-1}$ is well-defined as $\theta_i > 0$).

By Equation (2.114) and Assumption 2.6.5, we have

$$\lim_{N\to\infty}\frac{T}{N}\hat{C}\big(t(N)+\tfrac{T}{N},N\big) = \operatorname{diag}\big(\tfrac{f_{ii}(t)}{\theta_i}\big) = \tilde{F}(t)^{-1}$$

and therefore

$$\lim_{N\to\infty}\frac{N}{T}C\big(t(N)+\tfrac{T}{N},N\big)\hat{C}\big(t(N)+\tfrac{T}{N},N\big)^{-1}C\big(t(N)+\tfrac{T}{N},N\big) = F(t)\tilde{F}(t)F(t).$$

Furthermore,

$$\lim_{N\to\infty}\frac{N}{T}\Big(\Lambda(N)+D\big(t(N)+\tfrac{T}{N},N\big)\Big)^{-1} = \lim_{N\to\infty}\Big(\tfrac{N+1}{N}\Lambda+\tfrac{T}{N}D\big(t(N)+\tfrac{T}{N},N\big)\Big)^{-1} = \Lambda^{-1}$$

since

$$\lim_{N\to\infty}\hat{C}\big(t(N)+\tfrac{T}{N},N\big)^{-1} = 0.$$

We deduce

$$\lim_{N\to\infty}\tfrac{N}{T}\Big(C\big(t(N)+\tfrac{T}{N},N\big)-C\big(t(N),N\big)\Big) = -\alpha\Sigma + F(t)\tilde{F}(t)F(t) + F(t)\Lambda^{-1}F(t)$$

as desired. □

By Lemma 2.6.6, the limit of $(C(t(N),N))_N$ exists for a subsequence. There is reason to believe that all subsequences of $(C(t(N),N))_N$ converge to the same limit $F(t)$ and hence Assumption 2.6.5 holds in general. In this case Proposition 2.6.7 suggests that the limit F is differentiable and solves the Matrix Differentiable Equation (2.111) with boundary condition (2.112). As the principal solution for Riccati matrix differential equations is unique (cf. Remark 2.4.5), we believe that this is also the case for the Differential Equation (2.111) and hence that

$$F = C.$$

2.6.4. Convergence of the optimal liquidation strategy

Provided that Conjecture 2.6.2 holds, the convergence of the value function matrix carries over to convergence of the optimal strategy in the sense of Proposition 2.6.8 below. In particular, this yields a proof for convergence of the optimal strategy for the case $n = 1$.

We first introduce the following notation. For the discrete-time setting with $N+1$ trading times, we denote the optimal trading strategy at a given time $t = \frac{i(N)T}{N}$ for portfolio position $x \in \mathbb{R}^n$ at time t and the corresponding matrices A and B (cf. Equations (1.23) respectively (1.24)) by

$$x^*(t,N) = A(t,N)x,$$
$$y^*(t,N) = B(t,N)x.$$

We keep the notation for the continuous-time setting and denote the optimal trading strategy at time $t \in [0,T)$ and portfolio position x at time t by

$$u^*(t,x) = (\xi^*(t,x), \eta^*(t,x)).$$

By Theorem 1.3.4, the matrices $A(t,N)$ and $B(t,N)$ are given by

$$A(t,N) = \Big(\Lambda(N) + D\big(t+\tfrac{T}{N},N\big)\Big)^{-1} D\big(t+\tfrac{T}{N},N\big),$$
$$B(t,N) = \check{C}\big(t+\tfrac{T}{N},N\big)^{-1}\hat{P}(N)C\big(t+\tfrac{T}{N},N\big)\Big(I - A(t,N)\Big).$$

By Theorem 2.4.10, u^* is given by

$$\xi^*(t,x) = \Lambda^{-1}C(t)x,$$
$$\eta^*(t,x) = \bar{C}(t)C(t)x.$$

We obtain the following convergence result.

Proposition 2.6.8. *Assume that Conjecture 2.6.2 holds. Let $t \in [0,T)$ and $t(N) \in \{0, \frac{T}{N}, \ldots, T\}$ such that $\lim\limits_{N\to\infty} t(N) = t$. Let furthermore $x \in \mathbb{R}^n$ be the portfolio position at time t for the continuous-time setting respectively at time $t(N)$ for the discrete-time setting. Then*

$$\lim_{N\to\infty} \frac{x^*(t(N),N)}{T/N} = \xi^*(t,x)$$

and

$$\lim_{N\to\infty} y^*(t(N),N) = \eta^*(t,x).$$

Proof. Note first that Conjecture 2.6.2 implies

$$\lim_{N\to\infty} D\big(t(N)+\tfrac{T}{N},N\big) = C(t)$$

by Equation (2.114). Furthermore,

$$\lim_{N\to\infty} \frac{\Lambda(N) + D\big(t+\frac{T}{N},N\big)}{N} = \frac{\Lambda}{T}$$

and therefore

$$\begin{aligned}\lim_{N\to\infty} \frac{A(t(N),N)}{T/N} &= \frac{1}{T} \lim_{N\to\infty} N\Big(\Lambda(N) + D\big(t + \tfrac{T}{N}, N\big)\Big)^{-1} \lim_{N\to\infty} D\big(t(N) + \tfrac{T}{N}, N\big) \\ &= \Lambda^{-1} C(t)\end{aligned}$$

as required.

For optimal dark pool orders, we define the matrix $\bar{C}(t,N) = (\bar{c}_{i,j}(t,N))_{i,j=1,\ldots,n}$ by

$$\bar{C}(t,N) = \check{C}(t,N)^{-1} \hat{P}(N).$$

We obtain

$$\lim_{N\to\infty} \bar{C}\big(t(N) + \tfrac{t}{N}, N\big)^{-1} = \operatorname{diag}(c_{i,i}(t))$$

and therefore

$$\begin{aligned}\lim_{N\to\infty} B(t(N),N) &= \lim_{N\to\infty} \bar{C}\big(t + \tfrac{T}{N}, N\big)^{-1} C\big(t + \tfrac{T}{N}, N\big) \lim_{N\to\infty} \Big(I - A(t,N)\Big) \\ &= \bar{C}(t) C(t),\end{aligned}$$

finishing the proof. $\square$

3. Adverse selection in continuous time

In a similar fashion as in Chapter 2 we transfer the discrete-time market model of Section 1.4 into a continuous-time model and specify the costs of a trading strategy by the continuous-time analog of the the Optimization Problem ($\overline{\mathrm{OPT}_{\mathrm{dis}}}$). We set up the model in Section 3.1.

Recall that much of the difficulty in Chapter 2 stems from the fact that we consider n-dimensional portfolios. Here, we consider *single asset* liquidation. The fact that dark pool orders inhabit "linear" costs (in an analog way as in the discrete-time setting) gives rise to an optimization problem that is not linear-quadratic anymore. The value function is therefore not quadratic and the major difficulty is hence to derive the structure of the value function.

We derive a candidate for the value function in closed form heuristically in Section 3.2; this is accomplished by using the insights from the discrete-time setting of Section 1.4 and by analyzing the HJB equation corresponding to the problem. This candidate is a quadratic polynomial for large and for small asset positions. For intermediate positions, we "interpolate" these polynomials in a non-trivial way. We are also able to obtain a candidate for the optimal strategy in closed form.

It is a priori not clear that the candidate value function obtained in Section 3.2 is well-defined. We verify this in Section 3.3 and show that it is additionally continuously differentiable and strictly convex. We compute the partial derivatives which turn out to be of surprisingly simple form. Combining these results, we prove that the candidate value function solves the HJB equation corresponding to the problem, with the candidate optimal strategy as unique maximizer.

We finally solve the optimization problem rigorously in Section 3.4 via a verification based on the HJB equation. This involves taking the limit $s \to T-$ and requires preliminary considerations.

We close the chapter by studying the properties of this solution. In particular, we analyze the dependence of the value function and the optimal strategy on the extent of adverse selection and the optimal strategy for risk-neutral investors.

3.1. Model description

We consider the same model as in Chapter 2 but only treat the case of liquidating a *single asset position* ($n = 1$). In particular, the price impact is linear and temporary and is given by a positive real number Λ, the fundamental price process $\tilde{P}$ is a martingale and is only relevant through its variance $\Sigma \geq 0$, constant in time, and execution in the dark pool is driven by a one-dimensional Poisson process π with intensity $\theta \geq 0$. The set of admissible liquidation strategies is given as before by Definition 2.1.1; we only change the cost functional and add a term reflecting adverse selection. Therefore, trading in the dark pool is not entirely free anymore.

We define the cost functional and introduce the corresponding optimization problem in Section 3.1.1. In Section 3.1.2 we state the HJB equation for the optimization problem.

3.1.1. Cost functional

As in Section 1.4, we assume that there is adverse selection in the market. In the discrete-time setting the value function and the optimal strategy are independent of the fundamental asset price $\tilde{P}$, and we define the cost functional

$$\bar{J} : [0,T] \times \mathbb{R} \times \mathbb{A}(t,x) \longrightarrow \mathbb{R}_+$$

as the continuous-time analog of the cost functional of the Optimization Problem ($\overline{\mathrm{OPT}_{\mathrm{dis}}}$):

$$\bar{J}(t,x,u) := \mathbb{E}_{t,x}\Big[\int_t^T \bar{f}(\xi(s),\eta(s),X^u(s))ds\Big]; \tag{3.1}$$

the function $\bar{f} : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is given by

$$\bar{f}(\xi,\eta,x) := \Lambda\xi^2 + \theta\Gamma|\eta| + \alpha\Sigma x^2, \tag{3.2}$$

where $\Gamma \geq 0$ and $\theta \geq 0$ is the intensity of the Poisson process π. The value function of the optimization problem is

$$\bar{v}(t,x) := \inf_{u \in \mathbb{A}(t,x)} \bar{J}(t,x,u). \tag{$\overline{\mathrm{OPT}}$}$$

Thereby Γ denotes the expected price move in the primary exchange directly after liquidity in the dark pool is found. The special case $\Gamma = 0$ is included in Chapter 2 and solved explicitly in Section 2.5.1; $\theta = 0$ implies that the value of Γ is irrelevant. Therefore, we assume

$$\theta\Gamma > 0$$

from now on.

3.1.2. Hamilton-Jacobi-Bellman equation

The derivation of a candidate value function

$$w : [0,T) \times \mathbb{R} \longrightarrow \mathbb{R}_+$$

is the key step towards the solution of the Optimization Problem ($\overline{\mathrm{OPT}}$). Heuristic considerations suggest that w should satisfy the following HJB equation (cf. Section 2.1.3):

$$
\begin{aligned}
\frac{\partial w}{\partial t}(t,x) &= \sup_{u=(\xi,\eta)\in\mathbb{R}\times\mathbb{R}} \Big[\theta\big(w(t,x)-w(t,x-\eta)\big)+\frac{\partial w}{\partial x}(t,x)\xi-\bar{f}(\xi,\eta,x)\Big], \\
\lim_{t\to T-} w(t,x) &= \begin{cases} 0 & \text{if } x=0 \\ \infty & \text{else.} \end{cases}
\end{aligned}
\tag{$\overline{\text{HJB}}$}
$$

Because of the non-linear-quadratic form of the cost functional caused by the term $\theta\Gamma|\eta|$, there is no reason to believe that the value function is of linear-quadratic form, and a linear-quadratic ansatz as in Chapter 2 fails.

3.2. Heuristic derivation of the candidate value function

Given the shape of the cost functional and the results from the discrete-time setting in Section 1.4, we expect the value function and the optimal strategy to have the following structure:

- For large absolute values of the asset position $|x|$, the value function is a quadratic polynomial. The optimal strategy is affine linear in x.
- For small absolute values of the asset position $|x|$, the value function is given by

  $$w(t,x) = C(t;0)x^2,$$

  i.e., it is the same as the value function of the Optimization Problem (OPT) for $n=1$ and $\theta=0$ (cf. Equation (2.91)). Also the optimal strategy in the primary exchange is the same as the one without dark pool, and the optimal order in the dark pool is zero.
- For intermediate asset positions, we have to "interpolate" the value function in such a way that it is continuously differentiable.

In Section 3.2.1 we make the ansatz that the value function is a "quasi-polynomial" of degree two with the coefficients depending on the asset position x. In Section 3.2.2 we deduce a candidate for the optimal strategy (dependent on the coefficients of this polynomial) using the HJB-Equation ($\overline{\text{HJB}}$) and obtain a candidate for the boundary below which the dark pool is not used. In Section 3.2.3 we derive differential equations for the coefficients of the candidate value function.

3.2.1. Quasi-polynomial ansatz for the candidate value function

We make the ansatz that the value function is a quasi-polynomial of degree two with the coefficients depending on the asset position x:

$$w(t,x) = \bar{C}_1(t,x)x^2 + \bar{C}_2(t,x)|x| + \bar{C}_3(t,x) \tag{3.3}$$

for functions $\bar{C}_i : [0,T) \times \mathbb{R} \to \mathbb{R}$, $i = 1,2,3$. This ansatz is motivated by the fact that the value function is a piecewise quadratic polynomial in the discrete-time setting (cf. Theorem 1.4.4, in particular Equation (1.63)).

In the following we reflect on the optimal liquidation strategy. These considerations yield possible properties of the coefficients in Equation (3.3). Assume that

$$u^*(t,x) = (\xi^*(t,x), \eta^*(t,x))$$

is the optimal strategy. For small positions, the linear costs for trading in the dark pool are larger than the quadratic costs for trading in the primary venue. Therefore, there should be a time-dependent boundary $\beta : [0,T) \to \mathbb{R}_+$ such that

$$\eta^*(t,x) = 0 \quad \text{for } |x| \leq \beta(t). \tag{3.4}$$

For larger asset positions, the dark pool is cheaper than the primary venue and therefore we expect that the execution of the dark pool order decreases the position to the level β, where further dark pool use is too costly:

$$\eta^*(t,x) = \operatorname{sgn}(x)\big(|x| - \beta(t)\big) \quad \text{for } |x| > \beta(t). \tag{3.5}$$

As the dark pool is not used for $|x| \leq \beta(t)$, the optimal strategy below the boundary should be the one without dark pool, which we obtained in Chapter 2 (cf. Equation (2.91)), i.e, the value function is given by

$$w(t,x) = C(t;0)x^2, \tag{3.6}$$

and the optimal strategy in the primary venue is given by

$$\xi^*(t,x) = \frac{C(t;0)}{\Lambda} x.$$

In other words, we expect the coefficients of w as in Equation (3.3) to fulfill

$$\bar{C}_1(t,x) = C(t;0), \quad \bar{C}_2(t,x) = \bar{C}_3(t,x) = 0 \quad \text{for } |x| \leq \beta(t).$$

Let us now assume that these considerations are true. We hope that the value function is differentiable on $[0,T) \times \mathbb{R}$. We use this property for the proof of the verification theorem via the HJB Equation ($\overline{\text{HJB}}$) (cf. Theorem 3.4.4 below).

A necessary condition for continuity of w at $x = \beta(t)$ is

$$\bar{C}_1(t,\beta(t)) = C(t;0), \quad \bar{C}_2(t,\beta(t)) = \bar{C}_3(t,\beta(t)) = 0.$$

In order to be differentiable at $x = \beta(t)$, the right-hand and the left-hand partial derivative with respect to x must be equal, i.e., for $t \in [0,T)$, $x = \beta(t)$,

$$2C(t;0)x = 2\bar{C}_1(t,x)x + \bar{C}_2(t,x) + x^2\frac{\partial \bar{C}_1}{\partial x}(t,x) + x\frac{\partial \bar{C}_2}{\partial x}(t,x) + \frac{\partial \bar{C}_3}{\partial x}(t,x),$$

and therefore a necessary condition for differentiability is

$$\beta(t)^2\frac{\partial \bar{C}_1}{\partial x}(t,\beta(t))+\beta(t)\frac{\partial \bar{C}_2}{\partial x}(t,\beta(t))+\frac{\partial \bar{C}_3}{\partial x}(t,\beta(t))=0.$$

We make the educated guess that this condition also holds for $x \neq \beta(t)$:

$$x^2\frac{\partial \bar{C}_1}{\partial x}(t,x)+|x|\frac{\partial \bar{C}_2}{\partial x}(t,x)+\frac{\partial \bar{C}_3}{\partial x}(t,x)=0 \tag{3.7}$$

for all $t \in [0,T)$, $x \in \mathbb{R}$ and thus

$$\frac{\partial w}{\partial x}(t,x)=2\bar{C}_1(t,x)x+\operatorname{sgn}(x)\bar{C}_2(t,x). \tag{3.8}$$

3.2.2. The candidate optimal trading strategy

We assume from now on that the value function is given by w as in Equation (3.3) respectively as in Equation (3.6), that all partial derivatives exist, that $\frac{\partial w}{\partial x}$ is given as in Equation (3.8) and that the optimal strategy in the dark pool is given by Equations (3.4) and (3.5).

We consider the HJB Equation ($\overline{\text{HJB}}$) and maximize the function

$$\begin{aligned}h(t,x,\xi,\eta)&:=\theta\big(w(t,x)-w(t,x-\eta)\big)+\frac{\partial w}{\partial x}(t,x)\xi-\Lambda\xi^2-\theta\Gamma|\eta|-\alpha\Sigma x^2\\&=\theta w(t,x)-\theta\big(w(t,x-\eta)+\Gamma|\eta|\big)-\alpha\Sigma x^2\\&\quad-\Lambda\Big(\frac{\frac{\partial w}{\partial x}(t,x)}{2\Lambda}-\xi\Big)^2+\frac{\frac{\partial w}{\partial x}(t,x)^2}{4\Lambda}\end{aligned}$$

in (ξ,η). Using Equation (3.8), we obtain that h is maximal for

$$\xi^*:=\xi^*(t,x):=\frac{2\bar{C}_1(t,x)x+\operatorname{sgn}(x)\bar{C}_2(t,x)}{2\Lambda} \tag{3.9}$$

and for $\eta^*=\eta^*(t,x)$ such that

$$\bar{h}(\eta):=w(t,x-\eta)+\Gamma|\eta|$$

is minimal. For $x>\beta(t)$, Equation (3.5) suggests that this should be the case for

$$\eta^*=\eta^*(t,x)=x-\beta(t). \tag{3.10}$$

Furthermore, we have for $\eta^*>0$ and

$$\bar{h}'(\eta^*)=-\frac{\partial w}{\partial x}(t,x-\eta^*)+\Gamma=0$$

if and only if

$$\frac{\partial w}{\partial x}(t, x - \eta^*) = \Gamma. \tag{3.11}$$

Combining Equations (3.6), (3.10) and (3.11), we obtain the following candidate for the boundary $\beta(t)$:

$$\beta(t) = \frac{\Gamma}{2C(t;0)}. \tag{3.12}$$

3.2.3. Differential equations for the coefficients

From now on we consider the case $x > \beta(t)$. Given the assumption that w as in Equation (3.3) solves the HJB Equation ($\overline{\text{HJB}}$) with maximizer $u^* = (\xi^*, \eta^*)$ for ξ^* and η^* as in Equations (3.9) and (3.10), respectively, we obtain

$$\frac{\partial w}{\partial t}(t, x) = h(t, x, \xi^*, \eta^*).$$

Provided that Equations (3.7) and (3.12) hold, this implies

$$\begin{aligned}
&\frac{\partial \bar{C}_1}{\partial t}(t,x)x^2 + \frac{\partial \bar{C}_2}{\partial t}(t,x)x + \frac{\partial \bar{C}_3}{\partial t}(t,x) \\
&\quad = \theta\big(\bar{C}_1(t,x)x^2 + \bar{C}_2(t,x)x + \bar{C}_3(t,x)\big) - \theta C(t;0)\beta(t)^2 + \frac{1}{2\Lambda}\big(2\bar{C}_1(t,x)x + \bar{C}_2(t,x)\big)^2 \\
&\qquad - \frac{1}{4\Lambda}\big(2\bar{C}_1(t,x)x + \bar{C}_2(t,x)\big)^2 - \theta\Gamma(x - \beta(t))\big) - \alpha\Sigma x^2 \\
&\quad = \Big(\frac{\bar{C}_1(t,x)^2}{\Lambda} + \theta\bar{C}_1(t,x) - \alpha\Sigma\Big)x^2 + \Big(\bar{C}_2(t,x)\Big(\frac{\bar{C}_1(t,x)}{\Lambda} + \theta\Big) - \theta\Gamma\Big)x \\
&\qquad + \theta\bar{C}_3(t,x) + \frac{\theta\Gamma^2}{4C(t;0)} + \frac{\bar{C}_2(t,x)^2}{4\Lambda}.
\end{aligned}$$

We hence expect that for $|x| > \beta(t)$, the coefficients of the candidate value function fulfill the ordinary differential equations

$$\frac{\partial \bar{C}_1}{\partial t}(t,x) = \frac{\bar{C}_1(t,x)^2}{\Lambda} + \theta\bar{C}_1(t,x) - \alpha\Sigma, \tag{3.13}$$

$$\frac{\partial \bar{C}_2}{\partial t}(t,x) = \bar{C}_2(t,x)\Big(\frac{\bar{C}_1(t,x)}{\Lambda} + \theta\Big) - \theta\Gamma, \tag{3.14}$$

$$\frac{\partial \bar{C}_3}{\partial t}(t,x) = \theta\bar{C}_3(t,x) + \frac{\theta\Gamma^2}{4C(t;0)} + \frac{\bar{C}_2(t,x)^2}{4\Lambda} \tag{3.15}$$

and that for $x = \beta(t)$,

$$\bar{C}_1(t,x) = C(t;0), \quad \bar{C}_2(t,x) = \bar{C}_3(t,x) = 0. \tag{3.16}$$

In order to transfer the Differential Equations (3.13), (3.14), (3.15) and the Conditions (3.16) into initial value problems, we consider the trading trajectory resulting from the candidate optimal trading intensity ξ^*. Let $t \in [0,T)$, $x \geq \beta(t)$ and let $s \in [t,T)$ be the first time such that the trading trajectory crosses the boundary $\beta(s)$, provided that there is no dark pool execution in $[t,T)$. If such an s exists, it is the last time at which the optimal asset position crosses the boundary β, and we define

$$\bar{X}(t,s) := x.$$

In other words, the *deterministic* function $\bar{X}(\cdot,s)$ defined on $[t,s]$ is the candidate for the optimal asset position corresponding to the candidate optimal trading intensity ξ^*, provided no order in the dark pool is executed (cf. the respective functions $\bar{X}$ in the discrete-time setting of Section 1.4). Note also that this implies

$$\bar{X}(s,s) = \beta(s) = \frac{\Gamma}{2C(s;0)}.$$

We modify the notation for the coefficients of the candidate value function in the following way. If $t \in [0,T)$, $x \geq \beta(t)$ and if there exists a smallest $s \in [t,T)$ such that the trading trajectory resulting from the candidate optimal trading intensity ξ^* crosses the boundary β at time s without any dark pool execution, we write

$$C_1(t,s) := \bar{C}_1(t,x), \quad C_2(t,s) := \bar{C}_2(t,x), \quad C_3(t,s) := \bar{C}_3(t,x).$$

Thus, we expect that the new coefficients C_1, C_2 and C_3 solve the following initial value problems

$$\frac{\partial C_1}{\partial t}(\cdot,s) = \frac{C_1(\cdot,s)^2}{\Lambda} + \theta C_1(\cdot,s) - \alpha\Sigma \tag{3.17}$$
$$C_1(s,s) = C(s;0),$$
$$\frac{\partial C_2}{\partial t}(\cdot,s) = C_2(\cdot,s)\Big(\frac{C_1(\cdot,s)}{\Lambda} + \theta\Big) - \theta\Gamma \tag{3.18}$$
$$C_2(s,s) = 0,$$
$$\frac{\partial C_3}{\partial t}(\cdot,s) = \theta C_3(\cdot,s) + \frac{\theta\Gamma^2}{4C(\cdot;0)} + \frac{C_2(\cdot,s)^2}{4\Lambda} \tag{3.19}$$
$$C_3(s,s) = 0$$

and that $\bar{X}$ solves the initial value problem (cf. Equation (3.9))

$$\begin{aligned} \frac{\partial \bar{X}}{\partial t}(\cdot,s) &= -\frac{C_1(\cdot,s)}{\Lambda}\bar{X}(\cdot,s) - \frac{C_2(\cdot,s)}{2\Lambda} = -\xi^*(\cdot,\bar{X}(\cdot,s)) \\ \bar{X}(s,s) &= \frac{\Gamma}{2C(s;0)}. \end{aligned} \tag{3.20}$$

3.3. The candidate value function

Following the heuristic considerations of Section 3.2, we define the candidate value function w by

$$w(t,x) := C_1(t, g(t,x))x^2 + C_2(t, g(t,x))|x| + C_3(t, g(t,x)), \tag{3.21}$$

where C_1, C_2 and C_3 are the solutions of the Initial Value Problems (3.17), (3.18) and (3.19), respectively. The function $g : [0,T] \times \mathbb{R} \to [0,T]$ is given explicitly for $|x| \in [0, \frac{\Gamma}{2C(t;0)}] \cup [\bar{X}(t,T), \infty)$ by

$$g(t,x) := \begin{cases} t & \text{if } |x| \in [0, \frac{\Gamma}{2C(t;0)}] \\ T & \text{if } |x| \in [\bar{X}(t,T), \infty), \end{cases} \tag{3.22}$$

and implicitly for $|x| \in (\frac{\Gamma}{2C(t;0)}, \bar{X}(t,T))$ by

$$\bar{X}(t, g(t,x)) = |x|, \tag{3.23}$$

where $\bar{X}$ is the solution of the Initial Value Problem (3.20), i.e., $g(t,\cdot)$ is the inverse function of $\bar{X}(t,\cdot)$. It is not immediately clear that such an inverse function exists and that the function w defined by Equation (3.21) is well-defined.

The remainder of the section is structured as follows. In Section 3.3.1 we compute the solutions of the Initial Value Problems (3.17), (3.18), (3.19) and (3.20) in closed form. This enables us to prove that w is well-defined. In Section 3.3.2 we prove that w is continuously differentiable and strictly convex in x. Finally, we show that w is a solution of the HJB Equation ($\overline{\text{HJB}}$) in Section 3.3.3.

In Sections 3.3 - 3.5.1 we assume that

$$\alpha\Sigma > 0.$$

We treat the case $\alpha\Sigma = 0$ in Section 3.5.2.

3.3.1. Closed form solutions for the coefficients and the trading trajectory

in order that the value function is well-defined, the function g must be well-defined. Therefore, we require that $\bar{X}$ is strictly monotone in s (cf. Equation (3.23)). We prove this by directly computing the partial derivative of $\bar{X}$ with respect to s and show that it is strictly positive.

We start by giving closed form solutions for the value function coefficients C_1, C_2, C_3 and for $\bar{X}$, which follow from the Initial Value Problems (3.17), (3.18), (3.19) and (3.20), respectively. Proposition 3.3.1 treats the case $s \in (t,T)$, Proposition 3.3.2 treats the case $s = T$.

Proposition 3.3.1. *Let $t \in [0, T)$, $s \in (t, T)$ and define*

$$\begin{aligned}\mu(s) &:= \frac{2C(s;0) + \theta\Lambda}{\tilde{\theta}\Lambda}, \\ \kappa(s) &:= \operatorname{arcoth}(\mu(s));\end{aligned} \tag{3.24}$$

recall that

$$\tilde{\theta} = \sqrt{\theta^2 + \tfrac{4\alpha\Sigma}{\Lambda}}.$$

Then

$$C_1(t,s) = \frac{\Lambda\tilde{\theta}}{2}\coth\big(\frac{\tilde{\theta}}{2}(s-t) + \kappa(s)\big) - \frac{\Lambda\theta}{2} > 0, \tag{3.25}$$

$$\begin{aligned}C_2(t,s) = \frac{\theta\Gamma}{2\alpha\Sigma}\Big(&-\theta\Lambda \\ &+ \frac{\Lambda\tilde{\theta}(\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big) + \mu(s)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)) - 2C(s;0)\exp\big(\frac{\theta}{2}(t-s)\big)}{\mu(s)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big) + \cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)}\Big) \\ &> 0,\end{aligned} \tag{3.26}$$

$$C_3(t,s) = -\int_t^s \exp\big(\theta(t-u)\big)\Big(\frac{\theta\Gamma^2}{4C(u;0)} + \frac{C_2(u,s)^2}{4\Lambda}\Big)du < 0, \tag{3.27}$$

$$\begin{aligned}\bar{X}(t,s) = &\Big(\frac{\Gamma\mu(s)}{2C(s;0)} + \frac{\theta^2\Gamma}{2\tilde{\theta}\alpha\Sigma}\Big)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\theta}{2}(t-s)\big) \\ &+ \Big(\frac{\Gamma}{2C(s;0)} + \frac{\theta\Gamma}{2\alpha\Sigma}\Big)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\theta}{2}(t-s)\big) - \frac{\theta\Gamma}{2\alpha\Sigma}.\end{aligned} \tag{3.28}$$

Proof. Note first that for $s \in (t, T)$,

$$C(s;0) > \sqrt{\alpha\Sigma\Lambda}$$

by Equation (2.91). Therefore,

$$\Lambda\tilde{\theta} = \sqrt{\Lambda^2\theta^2 + 4\alpha\Sigma\Lambda} < \Lambda\theta + 2\sqrt{\alpha\Sigma\Lambda} \leq \Lambda\theta + 2C(s;0);$$

thus

$$\mu(s) > 1$$

and $\kappa(s)$ as in Equation (3.24) is well-defined for all $s \in [0, T)$.

Let us now consider the initial value problem

$$\begin{aligned}P' &= P^2 + \theta P - \frac{\alpha\Sigma}{\Lambda} \\ P(s) &= \frac{C(s;0)}{\Lambda}.\end{aligned}$$

By Section 2.2.2, the solution is given by

$$P(t) = \frac{\tilde{\theta}}{2} \coth\big(\frac{\tilde{\theta}}{2}(s-t) + \kappa(s)\big) - \frac{\theta}{2}.$$

Furthermore, ΛP solves the Initial Value Problem (3.17), finishing the proof of Equation (3.25).

Note now that $C_1(\cdot, s)$ is continuous on $[t, s)$. Thus, the solution of the inhomogeneous linear Initial Value Problem (3.18) is given by

$$C_2(t,s) = -\theta\Gamma \exp\Big(\int_s^t \big(\frac{C_1(u,s)}{\Lambda} + \theta\big)du\Big) \int_s^t \exp\Big(\int_u^s \big(\frac{C_1(v,s)}{\Lambda} + \theta\big)dv\Big)du > 0 \tag{3.29}$$

for $t < s$. Elementary but tedious algebraic manipulations show that

$$\begin{aligned} C_2(t,s) = \frac{\theta\Gamma}{2\alpha\Sigma}\Big(&- \theta\Lambda \\ &+ \frac{\Lambda\tilde{\theta}(\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big) + \mu(s)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)) - 2C(s;0)\exp\big(\frac{\theta}{2}(t-s)\big)}{\mu(s)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big) + \cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)}\Big) \end{aligned} \tag{3.30}$$

as required (cf. Equation (3.26)). A detailed proof of Equation (3.30) is provided in Appendix A.3.

The assertion that C_3 as in Equation (3.27) solves the inhomogeneous linear Initial Value Problem (3.19) follows directly as $C_2(\cdot, s)$ and $\frac{1}{C(\cdot;0)}$ are continuous.

The solution of the inhomogeneous linear Initial Value Problem (3.20) is given by

$$\begin{aligned} \bar{X}(t,s) &= \exp\Big(\int_t^s \big(\frac{C_1(u,s)}{\Lambda}\big)du\Big) \\ &\qquad \Big(\frac{\Gamma}{2C(s;0)} - \int_s^t \exp\Big(\int_s^u \big(\frac{C_1(v,s)}{\Lambda}\big)dv\Big)\frac{C_2(u,s)}{2\Lambda}du\Big) \end{aligned} \tag{3.31}$$

$$\begin{aligned} &= \Big(\frac{\Gamma\mu(s)}{2C(s;0)} + \frac{\theta^2\Gamma}{2\tilde{\theta}\alpha\Sigma}\Big)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\theta}{2}(t-s)\big) \\ &\quad + \Big(\frac{\Gamma}{2C(s;0)} + \frac{\theta\Gamma}{2\alpha\Sigma}\Big)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\theta}{2}(t-s)\big) - \frac{\theta\Gamma}{2\alpha\Sigma}. \end{aligned} \tag{3.32}$$

Again, a detailed proof of Equation (3.32) is provided in Appendix A.3. □

For $t \in [0, T)$, $C_1(t, s)$, $C_2(t, s)$, $C_3(t, s)$ and $\bar{X}(t, s)$ as in Equations (3.25), (3.26), (3.27) and (3.28) are not defined for $s = T$ as $C(T; 0)$ is not defined and $\lim_{s \to T-} C(s; 0) = \infty$. However, we require these objects for the definition of the candidate value function

w. We define $C_1(t,T)$, $C_2(t,T)$, $C_3(t,T)$ and $\bar{X}(t,T)$ by

$$C_1(t,T) := \lim_{s\to T-} C_1(t,s), \qquad C_2(t,T) := \lim_{s\to T-} C_2(t,s),$$
$$C_3(t,T) := \lim_{s\to T-} C_3(t,s), \qquad \bar{X}(t,T) := \lim_{s\to T-} \bar{X}(t,s).$$

The following proposition ensures that these limits exist and hence that $C_1(t,T)$, $C_2(t,T)$, $C_3(t,T)$ and $\bar{X}(t,T)$ are well-defined.

Proposition 3.3.2. *Let $t \in [0,T)$. Then*

$$C_1(t,T) = \frac{\Lambda\tilde{\theta}}{2}\coth\big(\frac{\tilde{\theta}}{2}(T-t)\big) - \frac{\Lambda\theta}{2} > 0, \tag{3.33}$$

$$C_2(t,T) = \frac{\theta\Gamma\Lambda}{2\alpha\Sigma}\Big(\tilde{\theta}\coth\big(\frac{\tilde{\theta}}{2}(T-t)\big) - \tilde{\theta}\frac{\exp\big(\frac{\theta}{2}(t-T)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(T-t)\big)} - \theta\Big) > 0, \tag{3.34}$$

$$C_3(t,T) = -\int_t^T \exp\big(\theta(t-u)\big)\Big(\frac{\theta\Gamma^2}{4C(u;0)} + \frac{C_2(u,T)^2}{4\Lambda}\Big)du < 0, \tag{3.35}$$

$$\begin{aligned}\bar{X}(t,T) = &\Big(\frac{\Gamma}{\tilde{\theta}\Lambda} + \frac{\theta^2\Gamma}{2\tilde{\theta}\alpha\Sigma}\Big)\sinh\big(\frac{\tilde{\theta}}{2}(T-t)\big)\exp\big(\frac{\theta}{2}(t-T)\big)\\ &+ \frac{\theta\Gamma}{2\alpha\Sigma}\cosh\big(\frac{\tilde{\theta}}{2}(T-t)\big)\exp\big(\frac{\theta}{2}(t-T)\big) - \frac{\theta\Gamma}{2\alpha\Sigma}.\end{aligned} \tag{3.36}$$

Proof. Note first that

$$\lim_{s\to T-} \kappa(s) = 0.$$

Therefore, Equation (3.33) follows directly from Equation (3.25).

By Equation (A.28) and the fact that

$$\lim_{s\to T-} \frac{C(s;0)}{\sqrt{\mu(s)^2-1}} = \lim_{s\to T-} \frac{C(s;0)}{\mu(s)} = \frac{\Lambda\tilde{\theta}}{2}, \tag{3.37}$$

Equation (3.34) follows.

Furthermore,

$$\begin{aligned}C_3(t,T) &= -\lim_{s\to T-}\int_t^T \mathbb{1}_{[t,s]}(u)\exp\big(\theta(t-u)\big)\Big(\frac{\theta\Gamma^2}{4C(u;0)} + \frac{C_2(u,s)^2}{4\Lambda}\Big)du\\ &= -\int_t^T \lim_{s\to T-}\Big(\mathbb{1}_{[t,s]}(u)\exp\big(\theta(t-u)\big)\Big(\frac{\theta\Gamma^2}{4C(u;0)} + \frac{C_2(u,s)^2}{4\Lambda}\Big)\Big)du\\ &= -\int_t^T \exp\big(\theta(t-u)\big)\Big(\frac{\theta\Gamma^2}{4C(u;0)} + \frac{C_2(u,T)^2}{4\Lambda}\Big)du\end{aligned} \tag{3.38}$$

(observe that $\frac{1}{C(s;0)} \to 0$ as $s \to T-$). We applied the dominated convergence theorem in Equation (3.38); recall that $\lim_{s\to T-} C(s;0) = \infty$ and that C_2 can be extended continuously to $A := \{(t,s)|t \in [0,T], s \in [t,T]\}$, hence it attains its maximum in A as $\lim_{t\to s-} C_2(t,s) = 0$ for all $s \in [0,T]$.

Equation (3.36) follows directly from Equation (3.28) and Equation (3.37). □

It follows immediately from the Formulae (3.25) - (3.28) that both the coefficients of the candidate value function and the trajectories $\bar{X}$ are continuously differentiable, even analytic.

Corollary 3.3.3. *Let $t \in [0,T)$, $s \in (t,T)$. Then C_1, C_2, C_3 and $\bar{X}$ have continuous partial derivatives of any order in (t,s).*

In order to show that the function g given in Equations (3.22) and (3.23) and thus the candidate value function is well-defined, we need monotonicity of the function $\bar{X}(t,\cdot)$. We prove this result by directly computing the partial derivative of $\bar{X}$ with respect to the second argument.

Proposition 3.3.4. *Let $t \in [0,T)$. Then $\bar{X}(t,\cdot)$ is strictly increasing in (t,T).*

Proof. Let $t \in [0,T)$, $s \in [t,T)$. Then

$$\frac{\partial \bar{X}}{\partial s}(t,s) = \frac{\alpha\Sigma\Gamma \exp\left(\frac{\theta}{2}(t-s)\right)}{2C(s;0)^2}\left(\mu(s)\sinh\left(\frac{\tilde{\theta}}{2}(s-t)\right) + \cosh\left(\frac{\tilde{\theta}}{2}(s-t)\right)\right) > 0 \tag{3.39}$$

since $\alpha\Sigma\Gamma > 0$ and $\mu(s) > 1$. A detailed proof of Equation (3.39) is provided in Appendix A.3. □

We can now prove the central result of this section.

Corollary 3.3.5. *The candidate value function w given by Equation* (3.21) *is well-defined.*

Proof. By Proposition 3.3.4, $\bar{X}(t,\cdot)$ possesses an inverse on (t,T). Thus, g as in Equations (3.22) and (3.23) is well-defined on $[0,T) \times \mathbb{R}$ (cf. also Equation (3.36)). Furthermore, $g(t,x) = g(t,|x|) \in [t,T]$ for all $t \in [0,T)$, $x \in \mathbb{R}$, and the functions C_i $(i = 1,2,3)$ are well-defined for $t \in [0,T)$, $s \in [t,T]$ by Propositions 3.3.1 and 3.3.2. □

3.3.2. Differentiation and convexity of the candidate value function

The goal of this section is to show that the candidate value function w is continuously differentiable. We are also able to compute the partial derivatives

$$\frac{\partial w}{\partial t}, \quad \frac{\partial w}{\partial x} \quad \text{and} \quad \frac{\partial^2 w}{\partial x^2}.$$

It turns out that $\frac{\partial w}{\partial x}$ has indeed the form we assumed in the heuristics in Section 3.2.1 (cf. Equation (3.8)). We also show that $\frac{\partial^2 w}{\partial x^2} > 0$ and hence that w is strictly convex in x. We first state the main theorem.

Theorem 3.3.6. *The candidate value function w given by Equation* (3.21) *is continuously differentiable in* $[0,T)\times\mathbb{R}$. *Furthermore, we have*

$$\frac{\partial w}{\partial x}(t,x) = 2C_1(t,g(t,x))x + \operatorname{sgn}(x)C_2(t,g(t,x)), \tag{3.40}$$

$$\frac{\partial w}{\partial t}(t,x) = \begin{cases} C'(t;0)x^2 & \text{if } x = \frac{\Gamma}{2C(t;0)} \\ \frac{\partial C_1}{\partial t}(t,g(t,x))x^2 + \frac{\partial C_2}{\partial t}(t,g(t,x))|x| + \frac{\partial C_3}{\partial t}(t,g(t,x)) & \text{else.} \end{cases} \tag{3.41}$$

Before we proceed with the proof of Theorem 3.3.6, we show that g is continuously differentiable. We prove this in the following lemma and compute the partial derivatives

$$\frac{\partial g}{\partial t} \quad \text{and} \quad \frac{\partial g}{\partial x}.$$

Lemma 3.3.7. *Let* $V := \{(t,x)\in\mathbb{R}^2 | t\in(0,T),\, x\in(\bar{X}(t,t),\bar{X}(t,T))\}$. *Then g given by Equations* (3.22) *and* (3.23) *is continuously differentiable on V and*

$$\frac{\partial g}{\partial x}(t,x) = \frac{1}{\frac{\partial \bar{X}}{\partial s}(t,g(t,x))}, \tag{3.42}$$

$$\frac{\partial g}{\partial t}(t,x) = -\frac{\frac{\partial \bar{X}}{\partial t}(t,g(t,x))}{\frac{\partial \bar{X}}{\partial s}(t,g(t,x))}. \tag{3.43}$$

Proof. Consider the open set $U := \{(t,s)\in\mathbb{R}^2 | t\in(0,T), s\in(t,T)\}$ and the continuously differentiable function $h:\mathbb{R}^2\longrightarrow\mathbb{R}^2$ (cf. Corollary 3.3.3),

$$h(t,s) := (t,\bar{X}(t,s))^\top.$$

Then for all $(t_0,s_0)\in U$, the Jacobian matrix of h is given by

$$Dh(t_0,s_0) := \begin{pmatrix} 1 & \frac{\partial \bar{X}}{\partial t}(t_0,s_0) \\ 0 & \frac{\partial \bar{X}}{\partial s}(t_0,s_0) \end{pmatrix}.$$

Note that $\frac{\partial \bar{X}}{\partial s}(t_0,s_0)\neq 0$ by Proposition 3.3.4. Therefore

$$\det Dh(t_0,s_0) = \frac{\partial \bar{X}}{\partial s}(t_0,s_0)\neq 0,$$

and we can apply the inverse function theorem. Thus, the inverse function h^{-1} of h exists locally and is locally continuously differentiable. However, on

$$\tilde{V} := h(U) = \{(t,x)\in\mathbb{R}^2 | t\in(0,T),\, x\in(\bar{X}(t,t),\bar{X}(t,T))\},$$

we have $h^{-1}=\tilde{g}$ for

$$\tilde{g}(t,x) := (t,g(t,x))^\top$$

globally, and therefore g is continuously differentiable on V.

Equation (3.42) follows form Equation (3.23) immediately. Equation (3.43) follows by differentiating with respect to t on either side of Equation (3.23). $\square$

We derive continuity of w directly.

Corollary 3.3.8. *w is continuous on $[0,T) \times \mathbb{R}$.*

We can now prove the main building block of the proof of Theorem 3.3.6. We show that

$$\frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s)^2 + \frac{\partial C_2}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_3}{\partial s}(t,s) = 0.$$

This is equivalent to Equation (3.7) in the heuristics of Section 3.2.1 and was the key step towards the derivation of the differential equations for the coefficients C_1, C_2 and C_3. Therefore, the construction of the candidate value function relies strongly on this property. Recall that we had good reasons to believe that Equation (3.7) is true for $x = \beta(t)$ or equivalently $t = s$. It is somewhat surprising that it is indeed also correct for $t < s$.

The second part of the lemma is a similar assertion which is useful for the computation of the second order derivative $\frac{\partial^2 w}{\partial x^2}$ and hence for the proof of the convexity of w below.

Lemma 3.3.9. *Let $t \in [0,T)$, and $s \in [t,T)$. Then*

(i)
$$a(t,s) := \frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s)^2 + \frac{\partial C_2}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_3}{\partial s}(t,s) = 0,$$

(ii)
$$b(t,s) := 2\frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_2}{\partial s}(t,s) = 0.$$

Proof. (i) For fixed s, we derive a linear homogeneous differential equation for $a(\cdot,s)$ on $[0,s)$ and show that

$$a(s,s) = \lim_{t \to s-} a(t,s) = 0.$$

To this end, we start by computing the derivatives of C_1, C_2 and C_3 with respect to s:

$$\frac{\partial C_1}{\partial s}(t,s) = \frac{-\theta C(s;0)}{\left(\mu(s)\sinh\left(\frac{\tilde{\theta}}{2}(s-t)\right) + \cosh\left(\frac{\tilde{\theta}}{2}(s-t)\right)\right)^2}, \tag{3.44}$$

$$\frac{\partial C_2}{\partial s}(t,s) = \frac{\theta\Gamma}{2\alpha\Sigma\left(\mu(s)\sinh\left(\frac{\tilde{\theta}}{2}(s-t)\right) + \cosh\left(\frac{\tilde{\theta}}{2}(s-t)\right)\right)^2} \cdot \Big(\frac{2\theta\alpha\Sigma + (\theta^2+\tilde{\theta}^2)C(s;0)}{\tilde{\theta}}\sinh\left(\frac{\tilde{\theta}}{2}(s-t)\right)\exp\left(\frac{\theta}{2}(t-s)\right) + (2\theta C(s;0) + 2\alpha\Sigma)\cosh\left(\frac{\tilde{\theta}}{2}(s-t)\right)\exp\left(\frac{\theta}{2}(t-s)\right) - 2\theta C(s;0)\Big), \tag{3.45}$$

$$\frac{\partial C_3}{\partial s}(t,s) = -\frac{\theta\Gamma^2}{4C(s;0)}\exp\big(\theta(t-s)\big) - \int_t^s \exp\big(\theta(t-u)\big)\frac{\partial}{\partial s}\Big(\frac{C_2(u,s)^2}{4\Lambda}\Big)du. \quad (3.46)$$

Detailed proofs of Equations (3.44) and (3.45) can be found in the Appendix A.3. Equation (3.46) follows from Leibniz' integral rule.

Using Equation (3.28), we deduce from Equations (3.44) - (3.46) that

$$\lim_{t\to s-} a(t,s) = -\theta C(s;0)\Big(\frac{\Gamma}{2C(s;0)}\Big)^2 + \theta\Gamma\frac{\Gamma}{2C(s;0)} - \frac{\theta\Gamma^2}{4C(s;0)} = 0. \quad (3.47)$$

Moreover, by Equation (3.46), we have

$$\begin{aligned} a(t,s) = {} & \frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s)^2 + \frac{\partial C_2}{\partial s}(t,s)\bar{X}(t,s) - \exp\big(\theta(t-s)\big)\frac{\theta\Gamma^2}{4C(s;0)} \\ & + \exp(\theta t)\int_t^s -\frac{\partial}{\partial s}\Big(\frac{C_2(u,s)^2}{4\Lambda}\Big)\exp(-\theta u)du. \end{aligned}$$

Thus, using the Differential Equations (3.17), (3.18), (3.19) and (3.20) (note that we can apply Schwarz' theorem by Corollary 3.3.3), we obtain

$$\begin{aligned} \frac{\partial a}{\partial t}(t,s) = {} & \frac{\partial^2 C_1}{\partial s\partial t}(t,s)\bar{X}(t,s)^2 + 2\frac{\partial C_1}{\partial s}(t,s)\frac{\partial \bar{X}}{\partial t}(t,s)\bar{X}(t,s) + \frac{\partial^2 C_2}{\partial s\partial t}(t,s)\bar{X}(t,s) \\ & + \frac{\partial C_2}{\partial s}(t,s)\frac{\partial \bar{X}}{\partial t}(t,s) - \theta\exp\big(\theta(t-s)\big)\frac{\theta\Gamma^2}{4C(s;0)} \\ & + \frac{\partial}{\partial s}\Big(\frac{C_2(t,s)^2}{4\Lambda}\Big) + \theta\exp(\theta t)\int_t^s -\frac{\partial}{\partial s}\Big(\frac{C_2(u,s)^2}{4\Lambda}\Big)\exp(-\theta u)du \\ = {} & \frac{\partial}{\partial s}\Big(\frac{C_1(t,s)^2}{\Lambda} + \theta C_1(t,s) - \alpha\Sigma\Big)\bar{X}(t,s)^2 \\ & + 2\frac{\partial C_1}{\partial s}(t,s)\Big(-\frac{C_1(t,s)}{\Lambda}\bar{X}(t,s) - \frac{C_2(t,s)}{2\Lambda}\Big)\bar{X}(t,s) \\ & + \frac{\partial}{\partial s}\Big(C_2(t,s)\Big(\frac{C_1(t,s)}{\Lambda} + \theta\Big) - \theta\Gamma\Big)\bar{X}(t,s) \\ & + \frac{\partial C_2}{\partial s}(t,s)\Big(-\frac{C_1(t,s)}{\Lambda}\bar{X}(t,s) - \frac{C_2(t,s)}{2\Lambda}\Big) \\ & - \theta\exp\big(\theta(t-s)\big)\frac{\theta\Gamma^2}{4C(s;0)} + \frac{\partial}{\partial s}\Big(\frac{C_2(t,s)^2}{4\Lambda}\Big) \\ & + \theta\exp(\theta t)\int_t^s -\frac{\partial}{\partial s}\Big(\frac{C_2(u,s)^2}{4\Lambda}\Big)\exp(-\theta u)du \\ = {} & \theta a(t,s). \end{aligned}$$

Hence, there exists a constant $C(s)$ $(s \in (0,T))$ such that for all $t \in [0,s)$,

$$a(t,s) = C(s)\exp(\theta t).$$

Equation (3.47) implies

$$0 = \lim_{t \to s-} a(t,s) = C(s)\exp(\theta s),$$

i.e., $C(s) = 0$, finishing the proof of (i).

(ii) We proceed similarly as in the proof of (i). Using the Differential Equations (3.17), (3.18) and (3.20), we obtain

$$\begin{aligned}\frac{\partial b}{\partial t}(t,s) &= 2\frac{\partial}{\partial s}\Big(\frac{C_1(t,s)^2}{\Lambda} + \theta C_1(t,s) - \alpha\Sigma\Big)\bar{X}(t,s)\\ &\quad + 2\frac{\partial C_1}{\partial s}(t,s)\Big(-\frac{C_1(t,s)}{\Lambda}\bar{X}(t,s) - \frac{C_2(t,s)}{2\Lambda}\Big)\\ &\quad + \frac{\partial}{\partial s}\Big(C_2(t,s)\Big(\frac{C_1(t,s)}{\Lambda} + \theta\Big) - \theta\Gamma\Big)\\ &= \Big(\frac{C_1(t,s)}{\Lambda} + \theta\Big)b(t,s).\end{aligned}$$

Thus, there exists a constant $C(s)$ $(s \in (0,T))$ such that for all $t \in [0,s)$,

$$b(t,s) = C(s)\exp\Big(\int_s^t \big(\frac{C_1(u,s)}{\Lambda} + \theta\big)du\Big).$$

Furthermore, by Equations (3.44) and (3.45) and Equation (3.28), we have

$$\lim_{t \to s-} b(t,s) = 0.$$

We deduce

$$0 = \lim_{t \to s-} b(t,s) = C(s),$$

finishing the proof.

□

We are now able to prove that w is continuously differentiable.

Proof of Theorem 3.3.6. Let $t \in [0,T), x \in \mathbb{R}$. We only treat the case $x \geq 0$. The case $x < 0$ follows accordingly. We consider four cases separately.

(i) $x < \frac{\Gamma}{2C(t;0)}$.
In this case we have $g(v,y) = v$ for all (v,y) in some neighborhood of (t,x) (cf. Equation (3.22)). Therefore,

$$w(v,y) = C(v;0)y^2$$

and consequently (recall that $C_1(t,t) = C(t;0)$, $C_2(t,t) = 0$)

$$\begin{aligned}\frac{\partial w}{\partial x}(t,x) &= 2C(t;0)x \\ &= 2C_1(t,t)x + C_2(t,t) \\ &= 2C_1(t,g(t,x))x + C_2(t,g(t,x)) \qquad (3.48)\end{aligned}$$

and

$$\frac{\partial w}{\partial t}(t,x) = C'(t;0)x^2. \qquad (3.49)$$

(ii) $x \in (\frac{\Gamma}{2C(t;0)}, \bar{X}(t,T))$.
This is equivalent to $x = \bar{X}(t,s)$, $g(t,x) = s$ for some $s \in (t,T)$. We have $y \in (\frac{\Gamma}{2C(v;0)}, \bar{X}(t,T))$ and therefore $g(v,y) \in (t,T)$ for all (v,y) in some neighborhood of (t,x), thus

$$w(t,y) = C_1(t,g(t,y))y^2 + C_2(t,g(t,y))y + C_3(t,g(t,y)).$$

By Lemma 3.3.7, we have

$$\frac{\partial g}{\partial x}(t,x) = \frac{1}{\frac{\partial \bar{X}}{\partial s}(t,g(t,x))} = \frac{1}{\frac{\partial \bar{X}}{\partial s}(t,s)} > 0$$

and hence by Lemma 3.3.9 (i),

$$\begin{aligned}\frac{\partial w}{\partial x}(t,x) &= 2C_1(t,g(t,x))x + C_2(t,g(t,x)) \\ &\quad + \frac{\partial g}{\partial x}(t,x)\Big(x^2\frac{\partial C_1}{\partial s}(t,g(t,x)) + x\frac{\partial C_2}{\partial s}(t,g(t,x)) + \frac{\partial C_3}{\partial s}(t,g(t,x))\Big) \\ &= 2C_1(t,g(t,x))x + C_2(t,g(t,x)) \\ &\quad + \frac{\partial g}{\partial x}(t,x)\Big(\frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s)^2 + \frac{\partial C_2}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_3}{\partial s}(t,s)\Big) \\ &= 2C_1(t,g(t,x))x + C_2(t,g(t,x)). \qquad (3.50)\end{aligned}$$

Similarly, we have

$$\frac{\partial g}{\partial t}(v,x) = -\frac{\frac{\partial \bar{X}}{\partial t}(v,g(v,x))}{\frac{\partial \bar{X}}{\partial s}(v,g(v,x))}$$

by Lemma 3.3.7 and therefore by Lemma 3.3.9 (i),

$$\begin{aligned}\frac{\partial w}{\partial t}(t,x) &= \frac{\partial C_1}{\partial t}(t,g(t,x))x^2 + \frac{\partial C_2}{\partial t}(t,g(t,x))x + \frac{\partial C_3}{\partial t}(t,g(t,x)) \\ &\quad + \frac{\partial g}{\partial t}(t,x)\Big(x^2\frac{\partial C_1}{\partial s}(t,g(t,x)) + x\frac{\partial C_2}{\partial s}(t,g(t,x)) + \frac{\partial C_3}{\partial s}(t,g(t,x))\Big)\end{aligned}$$

$$\begin{aligned}
&= \frac{\partial C_1}{\partial t}(t,g(t,x))x^2 + \frac{\partial C_2}{\partial t}(t,g(t,x))x + \frac{\partial C_3}{\partial t}(t,g(t,x)) \\
&\quad + \frac{\partial g}{\partial t}(t,x)\Big(\frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s)^2 + \frac{\partial C_2}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_3}{\partial s}(t,s)\Big) \\
&= \frac{\partial C_1}{\partial t}(t,g(t,x))x^2 + \frac{\partial C_2}{\partial t}(t,g(t,x))x + \frac{\partial C_3}{\partial t}(t,g(t,x)). \qquad (3.51)
\end{aligned}$$

(iii) $x > \bar{X}(t,T)$.
In this case we have $y > \bar{X}(v,T)$ and $g(v,y) = T$ for all (v,y) in some neighborhood of (t,x) (cf. Equation (3.22) again), thus

$$w(v,y) = C_1(v,T)y^2 + C_2(v,T)y + C_3(v,T).$$

Therefore,

$$\frac{\partial w}{\partial x}(t,x) = 2C_1(t,T)x + C_2(t,T) = 2C_1(t,g(t,x))x + C_2(t,g(t,x)), \qquad (3.52)$$

$$\frac{\partial w}{\partial t}(t,x) = \frac{\partial C_1}{\partial t}(t,T)x^2 + \frac{\partial C_2}{\partial t}(t,T)x + \frac{\partial C_3}{\partial t}(t,T). \qquad (3.53)$$

(iv) $x \in \{\frac{\Gamma}{2C(t;0)}, \bar{X}(t,T)\}$.
By Corollary 3.3.8, w is continuous. Furthermore, as $g(t,\cdot)$ is continuous in x with values in $[t,T]$ and C_1 and C_2 are continuous in $[t,T]$, $\frac{\partial w}{\partial x}(t,\cdot)$ can be extended continuously to $\mathbb{R}$ (cf. also Equations (3.48), (3.50) and (3.52)). Let now $h_n \neq 0$ ($n \in \mathbb{N}$), $h_n \to 0$. For large enough n, the mean-value theorem implies that there exists ξ_n strictly between x and $x - h_n$ such that

$$\frac{w(t,x) - w(t,x-h_n)}{h_n} = \frac{\partial w}{\partial x}(t,\xi_n).$$

As $\lim_{n\to\infty} \xi_n = x$, this implies

$$\frac{\partial w}{\partial x}(t,x) = \lim_{n\to\infty} \frac{w(t,x) - w(t,x-h_n)}{h_n} = \lim_{n\to\infty} \frac{\partial w}{\partial x}(t,\xi_n),$$

i.e., $w(t,\cdot)$ is differentiable in x with derivative

$$\frac{\partial w}{\partial x}(t,x) = \lim_{n\to\infty} \frac{\partial w}{\partial x}(t,\xi_n) = 2C_1(t,g(t,x))x + C_2(t,g(t,x)).$$

For differentiation with respect to t, note that by the differential equations for $C(t;0)$, $C_1(\cdot,s)$, $C_2(\cdot,s)$ and $C_3(\cdot,s)$, we obtain for $x = \frac{\Gamma}{2C(t;0)}$ that

$$\frac{\partial C_1}{\partial t}(t,t)x^2 + \frac{\partial C_2}{\partial t}(t,t)x + \frac{\partial C_3}{\partial t}(t,t) = C'(t;0)x^2.$$

Therefore, $\frac{\partial w}{\partial t}(\cdot,x)$ can be extended continuously to $[0,T)$ by Equations (3.49), (3.51) and (3.53). Using the mean-value theorem in a similar fashion as before, we obtain that $w(\cdot, x)$ is differentiable in t with derivative

$$\frac{\partial w}{\partial t}(t,x) = \begin{cases} \frac{\partial C_1}{\partial t}(t,t)x^2 + \frac{\partial C_2}{\partial t}(t,t)x + \frac{\partial C_3}{\partial t}(t,t) = C'(t;0)x^2 & \text{if } x = \frac{\Gamma}{2C(t;0)} \\ \frac{\partial C_1}{\partial t}(t,T)x^2 + \frac{\partial C_2}{\partial t}(t,T)x + \frac{\partial C_3}{\partial t}(t,T) & \text{if } x = \bar{X}(t,T). \end{cases}$$

□

By Theorem 3.3.6, we obtain first order partial derivatives of the candidate value function w. We can deduce strict convexity of w in x by computing the second order partial derivative

$$\frac{\partial^2 w}{\partial x^2}.$$

The key step towards this goal is Lemma 3.3.9 (ii).

Theorem 3.3.10. *The candidate value function w given by Equation* (3.21) *is strictly convex in $x \in \mathbb{R}$.*

Proof. Let $t \in [0,T)$ and $x \in \mathbb{R}$. We consider the case $x \geq 0$. The case $x < 0$ can be proven accordingly. Thus, $x = \bar{X}(t,s)$ for some $s \in [t,T)$ or $x \geq \bar{X}(t,T)$. In the first case Theorem 3.3.6, Lemma 3.3.7 and Lemma 3.3.9 (ii) yield

$$\begin{aligned} \frac{\partial^2 w}{\partial x^2}(t,x) &= 2C_1(t,g(t,x)) + \frac{\partial g}{\partial x}(t,x)\Big(2x\frac{\partial C_1}{\partial s}(t,g(t,x)) + \frac{\partial C_2}{\partial s}(t,g(t,x))\Big) \\ &= 2C_1(t,g(t,x)) + \frac{\partial g}{\partial x}(t,x)\Big(2\frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_2}{\partial s}(t,s)\Big) \\ &= 2C_1(t,g(t,x)). \end{aligned} \tag{3.54}$$

For $x > \bar{X}(t,T)$, we obtain

$$\frac{\partial^2 w}{\partial x^2}(t,x) = 2C_1(t,T). \tag{3.55}$$

By Equations (3.54) and (3.55), $\frac{\partial^2 w}{\partial x^2}(t,x)$ can be extended continuously to $x = \bar{X}(t,T)$ and we obtain

$$\frac{\partial^2 w}{\partial x^2}(t,\bar{X}(t,T)) = 2C_1(t,T)$$

by the same argument as in the proof of Theorem 3.3.6. Therefore (and by the respective considerations for $x < 0$), the second partial derivative of w with respect to x exists and

$$\frac{\partial^2 w}{\partial x^2}(t,x) > 0,$$

hence w is strictly convex in x. □

3.3.3. Hamilton-Jacobi-Bellman equation

The fact that w is differentiable and the formulae for the partial derivatives finally enable us to prove that w solves the HJB Equation ($\overline{\text{HJB}}$).

Theorem 3.3.11. *The candidate value function w given by Equation (3.21) fulfills the HJB Equation ($\overline{\text{HJB}}$) with unique maximizer $u^* = (\xi^*, \eta^*)$, where*

$$\xi^* = \xi^*(t,x) := \frac{2C_1(t,g(t,x))x + \operatorname{sgn}(x)C_2(t,g(t,x))}{2\Lambda}, \tag{3.56}$$

$$\eta^* = \eta^*(t,x) := \begin{cases} \operatorname{sgn}(x)\Big(|x| - \frac{\Gamma}{2C(t;0)}\Big) & \text{if } |x| > \frac{\Gamma}{2C(t;0)} \\ 0 & \text{if } |x| \le \frac{\Gamma}{2C(t;0)}. \end{cases} \tag{3.57}$$

Proof. The theorem follows directly from the subsequent Lemma 3.3.12 and Theorem 3.3.6 by using the Differential Equations (3.17), (3.18) and (3.19) for the case $|x| > \frac{\Gamma}{2C(t;0)}$ and the differential equation

$$C'(\cdot;0) = \frac{C(\cdot;0)^2}{\Lambda} - \alpha\Sigma$$

for the case $|x| \le \frac{\Gamma}{2C(t;0)}$. □

Lemma 3.3.12. *Let $h : [0,T) \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ be given by*

$$h(t,x,\xi,\eta) := \theta\big(w(t,x) - w(t,x-\eta)\big) + \frac{\partial w}{\partial x}(t,x)\xi - \bar{f}(\xi,\eta,x).$$

For fixed $t \in [0,T)$ and $x \in \mathbb{R}$, $h(t,x,\cdot,\cdot)$ attains its unique maximum at $u^* = (\xi^*, \eta^*)$ for ξ^* and η^* as in Equations (3.56) and (3.57), respectively.

Moreover,

$$h(t,x,\xi^*,\eta^*) = \begin{cases} \Big(\frac{C_1(t,g(t,x))^2}{\Lambda} + \theta C_1(t,g(t,x)) - \alpha\Sigma\Big)x^2 & \\ \quad + \Big(C_2(t,g(t,x))\Big(\frac{C_1(t,g(t,x))}{\Lambda} + \theta\Big) - \theta\Gamma\Big)|x| & \\ \quad + \theta C_3(t,g(t,x)) + \frac{\theta\Gamma^2}{4C(t;0)} + \frac{C_2(t,g(t,x))^2}{4\Lambda} & \text{if } |x| > \frac{\Gamma}{2C(t;0)} \\ \Big(\frac{C(t;0)^2}{\Lambda} - \alpha\Sigma\Big)x^2 & \text{if } |x| \le \frac{\Gamma}{2C(t;0)}. \end{cases} \tag{3.58}$$

Proof. Let $t \in [0,T)$ and $x \in \mathbb{R}$. We have

$$h(t,x,\xi,\eta) = \theta w(t,x) - \alpha\Sigma x^2 + h_1(t,x,\xi) + h_2(t,x,\eta)$$

for

$$\begin{aligned} h_1(t,x,\xi) &:= -\Lambda\xi^2 + \frac{\partial w}{\partial x}(t,x)\xi \\ &= -\Lambda\Big(\frac{\frac{\partial w}{\partial x}(t,x)}{2\Lambda} - \xi\Big)^2 + \frac{\frac{\partial w}{\partial x}(t,x)^2}{4\Lambda}, \end{aligned}$$

$$h_2(t,x,\eta) := -\theta w(t,x-\eta) - \theta\Gamma|\eta|.$$

As $\Lambda > 0$, $h_1(t,x,\cdot)$ is strictly concave in ξ. Furthermore, $h_2(t,x,\cdot)$ is strictly concave in η by Theorem 3.3.10. Thus, $h(t,x,\cdot,\cdot)$ is strictly concave in (ξ,η) and attains its unique global maximum in (ξ^*,η^*) if $h_1(t,x,\cdot)$ attains its unique global maximum in ξ^* and h_2 attains its unique global maximum in η^*.

By Theorem 3.3.6, h_1 attains its unique global maximum for ξ^* as in Equation (3.56).

Note now that $h_2 \leq 0$. For $x = 0$,

$$h_2(t,x,\eta^*) = h_2(t,0,0) = 0,$$

and therefore h_2 attains its unique global maximum in η^* as in Equation (3.57).

For $\eta \neq 0$, h_2 is differentiable in η. For $|x| > \frac{\Gamma}{2C(t;0)}$, we have

$$\begin{aligned}\frac{\partial h_2}{\partial \eta}(t,x,\eta^*) &= \theta\Big(\frac{w}{\partial x}(t,x-\eta^*) - \operatorname{sgn}(x)\Gamma\Big)\\ &= \theta\Big(2C(t;0)\operatorname{sgn}(x)\frac{\Gamma}{2C(t;0)} - \operatorname{sgn}(x)\Gamma\Big)\\ &= 0\end{aligned}$$

for $\eta^* = \eta^*(t,x)$ as in Equation (3.57) (note that $\eta^*(t,x) \neq 0$ for $|x| > \frac{\Gamma}{2C(t;0)}$). By strict concavity, we conclude that $h_2(t,x,\cdot)$ attains its unique maximum at η^* as in Equation (3.57).

Let now $x \in (0, \frac{\Gamma}{2C(t;0)})$. We have

$$\begin{aligned}\lim_{\eta\to 0-}\frac{\partial h_2}{\partial \eta}(t,x,\eta) &= \lim_{\eta\to 0-} 2\theta C(t;0)(x-\eta) + \theta\Gamma = 2\theta C(t;0)x + \theta\Gamma > 0,\\ \lim_{\eta\to 0+}\frac{\partial h_2}{\partial \eta}(t,x,\eta) &= \lim_{\eta\to 0+} 2\theta C(t;0)(x-\eta) - \theta\Gamma = 2\theta C(t;0)x - \theta\Gamma < 0.\end{aligned}$$

Therefore, $\eta^* = 0$ maximizes h_2 uniquely by strict concavity as required.

By continuity of η^* and h_2, $\eta^*(t,x) = 0$ maximizes h_2 uniquely for $x = \frac{\Gamma}{2C(t;0)}$.

The case $x \in [-\frac{\Gamma}{2C(t;0)}, 0)$ follows similarly as above. We plug (ξ^*,η^*) into h and obtain Equation (3.58), finishing the proof. □

3.4. Optimal liquidation

In the previous sections we collected the building blocks for solving the Optimization Problem ($\overline{\text{OPT}}$). The goal of this section is to prove that ($\overline{\text{OPT}}$) is indeed solved by the candidate optimal strategy $u^* = (\xi^*,\eta^*)$ and that the value function is given by the candidate value function w.

Given the candidate optimal strategy u^*, we denote the process controlled by u^* by

$$X^* := X^{u^*}.$$

We proceed as follows. In Section 3.4.1 we obtain bounds for w, u^* and X^*. Using these bounds, we show that u^* is an admissible liquidation strategy. We prove the verification theorem in Section 3.4.2.

3.4.1. Admissibility of the candidate optimal strategy

By Equation (3.44), $C_1(t,\cdot)$ is strictly decreasing in $[t,T]$. Using the series expansion of coth, Equation (2.91) and Equation (3.33) we obtain the following inequalities (cf. also Equations (2.90) and (3.33)). For $t \in [0,T)$ and $s_1, s_2 \in (t,T)$, $s_1 < s_2$,

$$0 < C(t;\theta) = C_1(t,T) < C_1(t,s_2) < C_1(t,s_1) < C_1(t,t) = C(t;0) \leq \frac{\Lambda}{T-t} + \sqrt{\alpha\Sigma\Lambda}. \quad (3.59)$$

Using Taylor's theorem with integral remainder term, we can deduce the following useful bounds for the candidate value function w and the candidate optimal strategy in the primary venue ξ^*.

Proposition 3.4.1. *Let $t \in [0,T)$ and $x \neq 0$. Then*

$$C_1(t,T)x^2 < w(t,x) \leq C(t;0)x^2 \leq \Big(\frac{\Lambda}{T-t} + \sqrt{\alpha\Sigma\Lambda}\Big)x^2, \quad (3.60)$$

$$\frac{C_1(t,T)}{\Lambda}|x| < |\xi^*(t,x)| \leq \frac{C(t;0)}{\Lambda}|x| \leq \Big(\frac{1}{T-t} + \sqrt{\frac{\alpha\Sigma}{\Lambda}}\Big)|x|. \quad (3.61)$$

Proof. We use Taylor's theorem with integral remainder term. Note first that

$$w(t,0) = \frac{\partial w}{\partial x}(t,0) = 0$$

by Equations (3.21) and (3.40) (recall that $g(t,0) = t$). Thus by Equation (3.54),

$$\begin{aligned} w(t,x) &= w(t,0) + \frac{\partial w}{\partial x}(t,0)x + \int_0^x \frac{\partial^2 w}{\partial x^2}(t,y)(x-y)dy \\ &= 2\int_0^x C_1(t,g(t,y))(x-y)dy. \end{aligned}$$

Using (3.59), we can directly deduce (3.60). Furthermore,

$$\frac{\partial w}{\partial x}(t,x) = 2\Lambda\xi^*(t,x).$$

Thus,

$$2\Lambda\xi^*(t,x) = \frac{\partial w}{\partial x}(t,0) + \int_0^x \frac{\partial^2 w}{\partial x^2}(t,y)dy$$

$$= 2 \int_0^x C_1(t, g(t, y))dy,$$

and (3.61) follows from (3.59). □

Using Gronwall's inequality, we can deduce an upper bound for the candidate optimal asset position X^*.

Corollary 3.4.2. *Let $t \in [0, T)$ and $x \in \mathbb{R}$. Then*

$$|X^*(s)| \leq |x| \exp\Big(- \int_t^s \frac{C_1(u, T)}{\Lambda} du\Big) = |x| \exp\big(\frac{\Lambda\theta}{2}(s-t)\big) \frac{\sinh\big(\frac{\tilde{\theta}}{2}(T-s)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(T-t)\big)}.$$

Proof. We define the process $(\tilde{X}(s))_{s\in[t,T)}$ as the solution of the initial value problem

$$X'(s) = -\xi^*(s, X(s))$$
$$X(t) = x.$$

By the structure of the process η^* (cf. Equation (3.57)), we have that for all $s \in [0, T)$,

$$|X^*(s)| \leq |\tilde{X}(s)|.$$

Therefore, by Gronwall's inequality, Proposition 3.4.1 and Equation (3.33),

$$|X^*(s)| \leq |\tilde{X}(s)| \leq |x| \exp\Big(- \int_t^s \frac{C_1(u, T)}{\Lambda} du\Big) = |x| \exp\big(\frac{\Lambda\theta}{2}(s-t)\big) \frac{\sinh\big(\frac{\tilde{\theta}}{2}(T-s)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(T-t)\big)}$$

as required. □

The bounds derived above allow us to deduce admissibility of the candidate optimal strategy u^*.

Proposition 3.4.3. *Let $t \in [0, T)$, $x \in \mathbb{R}$ be the asset position at time t and $u^* = (\xi^*, \eta^*)$ be as in Equations (3.56) and (3.57). Then $u^* \in \mathbb{A}(t, x)$.*

Proof. Note first that (i) of Definition 2.1.1 is straightforward and that (iii) holds as $\theta > 0$. By Proposition 3.4.1, Corollary 3.4.2 and Equation (2.91) we have

$$\begin{aligned}
|\xi^*(s, X^*(s))| &\leq \frac{C(s; 0)}{\Lambda} |X^*(s)| \\
&\leq |x| \sqrt{\frac{\alpha\Sigma}{\Lambda}} \exp\big(\frac{\Lambda\theta}{2}(s-t)\big) \frac{\sinh\big(\frac{\tilde{\theta}}{2}(T-s)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(T-t)\big)} \frac{\cosh\big(\sqrt{\frac{\alpha\Sigma}{\Lambda}}(T-s)\big)}{\sinh\big(\sqrt{\frac{\alpha\Sigma}{\Lambda}}(T-s)\big)} \\
&\longrightarrow |x| \sqrt{\frac{\alpha\Sigma}{\Lambda}} \exp\big(\frac{\Lambda\theta}{2}(T-t)\big) \frac{1}{\sinh\big(\frac{\tilde{\theta}}{2}(T-t)\big)} < \infty
\end{aligned}$$

as $s \to T-$. Thus,

$$\mathbb{E}_{t,x}\Big[\int_t^T |\xi^*(s, X^*(s))|^4 ds\Big] < \infty.$$

As $|\eta^*(t, X^*(s-))| < |X^*(s-)|$, we have

$$\mathbb{E}_{t,x}\Big[\int_t^T |\eta^*(s, X^*(s-))|^8 ds\Big] < \infty$$

by Corollary 3.4.2. Finally, again by Corollary 3.4.2,

$$\lim_{s\to T-} X^*(s) = 0.$$

$\square$

3.4.2. Verification

We are now finally able to prove that the Optimization Problem ($\overline{\text{OPT}}$) is solved by the candidate optimal strategy in Equations (3.56) and (3.57) and that the value function is given as in Equation (3.21).

Theorem 3.4.4. *The value function of the Optimization Problem* ($\overline{\text{OPT}}$) *is given by*

$$\bar{v}(t, x) = C_1(t, g(t, x))x^2 + C_2(t, g(t, x))|x| + C_3(t, g(t, x))$$

for $t \in [0, T)$, $x \in \mathbb{R}^n$ (cf. Equation (3.21)*) where C_1, C_2, C_3 and g are given as in Section 3.3. The $\mathbb{P}\otimes\lambda$ - a.s. unique optimal liquidation strategy is given by $u^* = (\xi^*, \eta^*)$ such that (cf. Equations* (3.56) *and* (3.57)*)*

$$\xi^* = \xi^*(t, x) := \frac{2C_1(t, g(t, x))x + \operatorname{sgn}(x)C_2(t, g(t, x))}{2\Lambda}, \tag{3.62}$$

$$\eta^* = \eta^*(t, x) := \begin{cases} \operatorname{sgn}(x)\Big(|x| - \frac{\Gamma}{2C(t;0)}\Big) & \text{if } |x| > \frac{\Gamma}{2C(t;0)} \\ 0 & \text{if } |x| \le \frac{\Gamma}{2C(t;0)}. \end{cases} \tag{3.63}$$

For the proof of the theorem, we require the following lemma, which relies on the fact that the value function v of the Optimization Problem (OPT) is a lower bound of $\bar{v}$.

Lemma 3.4.5. *Let $t \in [0, T)$, $x \in \mathbb{R}$ and $u \in \mathbb{A}(t, x)$. Then*

$$\lim_{s\to T-} \mathbb{E}_{t,x}\big[w(s, X^u(s))\big] = 0$$

Proof. Note first that there exists a constant $K > 0$ such that

$$\frac{K}{T-t}x^2 \le C(t;\theta)x^2 = v(t, x) \le \bar{v}(t, x), \tag{3.64}$$

where v is the value function of the Optimization Problem (OPT) for optimal liquidation without adverse selection and $C(t;\theta)$ is as in Equation (2.90).

A strategy $u \in \mathbb{A}(t,x)$ has finite costs $\bar{J}(t,x,u) < \infty$ by Definition 2.1.1 (ii) and Proposition 2.1.4. Let now $s \in (t,T)$. Then

$$\begin{aligned}
\bar{J}(t,x,u) &= \mathbb{E}_{t,x}\Big[\int_t^s \Big(\Lambda\xi(r)^2 + \theta\Gamma|\eta(r)| + \alpha\Sigma X^u(r)^2\Big)dr\Big] \\
&\quad + \mathbb{E}_{t,x}\Big[\int_s^T \Big(\Lambda\xi(r)^2 + \theta\Gamma|\eta(r)| + \alpha\Sigma X^u(r)^2\Big)dr\Big] \\
&\geq \mathbb{E}_{t,x}\Big[\int_t^s \Big(\Lambda\xi(r)^2 + \theta\Gamma|\eta(r)| + \alpha\Sigma X^u(r)^2\Big)dr + \bar{v}(s, X^u(s))\Big] \\
&\overset{(3.64)}{\geq} \mathbb{E}_{t,x}\Big[\int_t^s \Big(\Lambda\xi(r)^2 + \theta\Gamma|\eta(r)| + \alpha\Sigma X^u(r)^2\Big)dr\Big] + K\mathbb{E}_{t,x}\Big[\frac{X^u(s)^2}{T-s}\Big].
\end{aligned}$$

Thus by the monotone convergence theorem,

$$\begin{aligned}
\bar{J}(t,x,u) &\geq \lim_{s\to T-}\Big(\mathbb{E}_{t,x}\Big[\int_t^s \Big(\Lambda\xi(r)^2 + \theta\Gamma|\eta(r)| + \alpha\Sigma X^u(r)^2\Big)dr\Big] + K\mathbb{E}_{t,x}\Big[\frac{X^u(s)^2}{T-s}\Big]\Big) \\
&= \bar{J}(t,x,u) + K\lim_{s\to T-}\mathbb{E}_{t,x}\Big[\frac{X^u(s)^2}{T-s}\Big],
\end{aligned}$$

hence

$$\lim_{s\to T-}\mathbb{E}_{t,x}\Big[\frac{X^u(s)^2}{T-s}\Big] = 0.$$

The assertion now follows directly from Proposition 3.4.1 (cf. Inequality (3.60)). □

Proof of Theorem 3.4.4. Let $t \in [0,T)$, $x \in \mathbb{R}$, $u \in \mathbb{A}(t,x)$ and $s \in [t,T)$. We apply Itô's formula to the function $w(r, X^u(r))$ (recall that w is differentiable by Theorem 3.3.6):

$$\begin{aligned}
w(t,x) &= w(s, X^u(s)) + \int_t^s \Big(\frac{\partial w}{\partial x}(r, X^u(r))\xi(r) - \frac{\partial w}{\partial r}(r, X^u(r))\Big)dr \\
&\quad + \int_t^s (w(r, X^u(r-)) - w(r, X^u(r-) - \eta(r)))\pi(dr) \\
&\leq w(s, X^u(s)) + \int_t^s \bar{f}(\xi(r), \eta(r), X^u(r))dr \\
&\quad + \int_t^s (w(r, X^u(r-)) - w(r, X^u(r-) - \eta(r)))\pi(dr)
\end{aligned}$$

$$- \int_t^s \theta\big(w(r, X^u(r-)) - w(r, X^u(r-) - \eta(r))\big) dr \tag{3.65}$$

$$= w(s, X^u(s)) + \int_t^s \bar{f}(\xi(r), \eta(r), X^u(r)) dr$$

$$+ \int_t^s \big(w(r, X^u(r-)) - w(l, r, X^u(r-) - \eta(r))\big) M(dr), \tag{3.66}$$

where M is the compensated Poisson process given by $M(r) := \pi(r) - \theta r$ and Inequality (3.65) follows from Theorem 3.3.11, with equality if and only if

$$\lambda[u = u^*] = 0.$$

Taking expectations on both sides, we obtain

$$w(t, x) \leq \mathbb{E}_{t,x}[w(s, X^u(s))] + \mathbb{E}_{t,x}\Big[\int_t^s \bar{f}(\xi(r), \eta(r), X^u(r)) dr\Big]$$

$$+ \mathbb{E}_{t,x}\Big[\int_t^s w(r, X^u(r-)) - w(r, X^u(r-) - \eta(r)) M(dr)\Big], \tag{3.67}$$

with equality if and only if

$$\lambda \otimes \mathbb{P}[u = u^*] = 0.$$

By Proposition 3.4.1, there exists a constant $K = K(s)$ such that

$$\mathbb{E}_{t,x}\Big[\int_t^s |w(r, X^u(r-)) - w(r, X^u(r-) - \eta(r, X^u(r-)))|^2 dr\Big]$$

$$\leq K(s)\Big(\mathbb{E}_{t,x}\Big[\int_t^s (|X^u(r-)|^4 + |\eta(r)|^4) dr\Big]\Big)$$

$$< \infty$$

by Definition 2.1.1 (ii) and Proposition 2.1.4. As $\langle M \rangle(s) = \theta s$, this implies that

$$\tilde{M}(s) := \int_t^s \big(w(r, X^u(r-)) - w(r, X^u(r-) - \eta(r))\big) M(dr)$$

is a martingale. Thus, taking the limit $s \to T-$ in Inequality (3.67), we obtain by

Lemma 3.4.5 and the monotone convergence theorem that

$$w(t,x) \leq \mathbb{E}_{t,x}\Big[\int_t^T \bar{f}(\xi(r),\eta(r),X^u(r))dr\Big] = \bar{J}(t,x,u), \tag{3.68}$$

again with equality if

$$\lambda \otimes \mathbb{P}[u = u^*] = 0.$$

Uniqueness follows from the strict convexity of $\bar{f}$ as in the proof of Theorem 2.4.10. □

3.5. Properties of the value function and the optimal strategy

In the discrete-time setting of Section 1.4 we discussed the properties of the solution of the discrete-time optimization problem corresponding to ($\overline{\text{OPT}}$) in detail (cf. also the heuristics in Section 3.2). Most of these properties hold in an analog way in the continuous-time setting. Therefore, the discussion will be shorter here.

In Section 3.5.1 we discuss the dependence of the optimal strategy and the value function on the adverse selection parameter Γ. In Section 3.5.2 we analyze the case of a risk-neutral investor ($\alpha\Sigma = 0$), which we had excluded in Sections 3.3 and 3.4.

3.5.1. Risk-averse investors

Theorem 3.4.4 confirms the structure of the optimal strategy and the value function of the Optimization Problem ($\overline{\text{OPT}}$) we had expected in the heuristics of Section 3.2.

Both for large initial asset positions x ($|x| \geq \bar{X}(t,T)$) and for small initial asset positions ($|x| \leq \bar{X}(t,t) = \frac{\Gamma}{2C(t;0)}$), the value function is a quadratic polynomial. In between, it is an "interpolation" of these polynomials.

The value function and the optimal strategy for $|x| \leq \frac{\Gamma}{2C(t;0)}$ are the same as the ones without dark pool and without adverse selection (i.e., the optimal order in the dark pool is zero).

For larger asset positions, the absolute value of order in the dark pool is greater than zero; after the execution of the optimal dark pool order at time s, the asset position is $\frac{\Gamma}{2C(s;0)}$. The optimal trading trajectory until dark pool execution for an initial asset positions $|x| = \bar{X}(t,s) \in (\bar{X}(t,t), \bar{X}(t,T))$ is given by the function $\bar{X}(\cdot, s)$ in $[t,s]$.

It is interesting to examine the dependence of the value function and the optimal strategy on the adverse selection parameter Γ. Intuitively, costs should be higher for large adverse selection and therefore the value function should be increasing in Γ. Furthermore, the dark pool is less attractive for large adverse selection and therefore optimal dark pool orders should be decreasing in Γ, whereas trading intensity in the primary venue should be increasing in Γ (as impact costs are relatively lower with respect to adverse selection costs). The following proposition confirms these intuitions. To stress the dependence on Γ, we will add it as an argument for the remainder of the section and write

$$\bar{v}(t,x,\Gamma) := \bar{v}(t,x), \quad u^*(t,x,\Gamma) := u^*(t,x), \quad \cdots$$

Proposition 3.5.1. *Let $t \in [0,T)$, $s \in [t,T]$ and $x \in \mathbb{R}$. Then*

(i) $C_1(t,s,\cdot)$ is constant, $C_2(t,s,\cdot)$ is strictly increasing and $C_3(t,s,\cdot)$ is strictly decreasing.

(ii) $\bar{X}(t,s,\cdot)$ is strictly increasing.

(iii) $\bar{v}(t,x,\cdot)$ is strictly increasing on the interval $(0, 2|x|C(t;0))$ and constant for $\Gamma > 2|x|C(t;0)$.

(iv) $|\xi^*(t,x,\cdot)|$ is strictly increasing and $|\eta^*(t,x,\cdot)|$ is strictly decreasing on the interval $(0, 2|x|C(t;0))$ and constant for $\Gamma > 2|x|C(t;0)$.

Proof. (i) The first assertion follows directly from the Initial Value Problem for C_1, (3.17). The second and the third assertion can be deduced from that by Equations (3.29) and (3.27), respectively.

(ii) The assertion follows directly from Equation (3.31).

(iii) Monotonicity of the value function follows directly from the form of the cost functional (cf. Equation (3.1)) as long as $|\eta^*(t,x,\Gamma)| > 0$, i.e., $\Gamma \in (0, 2|x|C(t;0))$ (cf. Equation (3.63)).

(iv) For the first assertion, let without loss of generality $x > 0$ and $\Gamma < \tilde{\Gamma} < 2xC(t;0)$. By (ii) (cf. Equation (3.23)),

$$g(t,y,\tilde{\Gamma}) \leq g(t,y,\Gamma)$$

for all $y \in [0,x]$ with strict inequality for

$$\frac{\tilde{\Gamma}}{2C(t;0)} < y < \bar{X}(t,T,\Gamma).$$

Therefore (cf. the proof of Proposition 3.4.1),

$$\xi^*(t,x,\Gamma) = \frac{1}{\Lambda}\int_0^x C_1(t,g(t,y,\Gamma),\Gamma)dy < \frac{1}{\Lambda}\int_0^x C_1(t,g(t,y,\tilde{\Gamma}),\tilde{\Gamma})dy = \xi^*(t,x,\tilde{\Gamma})$$

by (i) and the fact that C_1 is strictly decreasing in s (cf. Equation (3.44)).

The assertion that $|\eta^*(t,x,\cdot)|$ is strictly decreasing follows directly from Equation (3.63).

□

3.5.2. Risk-neutral investors

In Sections 3.3 and 3.4 we excluded the case $\alpha\Sigma = 0$. The main reason was that in some formulae the term $\alpha\Sigma$ appears in the denominator (cf., e.g., Equations (3.28), and (3.36)) and it would thus have complicated the exposition of the results. However, the case

$$\alpha\Sigma = 0$$

can be treated in a similar way as the case $\alpha\Sigma > 0$.

We define C_1, C_2, C_3 and $\bar{X}$ as before by the Initial Value Problems (3.17) - (3.20). Note that only the differential equation for C_1 depends on $\alpha\Sigma$. Therefore C_2, C_3 and $\bar{X}$ only depend on $\alpha\Sigma$ through C_1. We can compute solutions of the initial value problems directly or by taking the limits for $\alpha\Sigma \to 0$ in Equations (3.25) - (3.28).

Note first that $\alpha\Sigma = 0$ implies

$$\tilde{\theta} = \theta, \quad C(t;0) = \frac{\Lambda}{T-t}.$$

Therefore,

$$\bar{X}(t,t) = \frac{\Gamma}{2C(t;0)} = \frac{\Gamma}{2\Lambda}(T-t).$$

By Proposition 3.3.4, we have

$$\frac{\partial \bar{X}}{\partial s}(t,s) = 0$$

and therefore

$$\bar{X}(t,s) = \frac{\Gamma}{2\Lambda}(T-t) \tag{3.69}$$

for $t \in [0,T)$, $s \in [t,T]$. Hence, we expect only two different trading regions. The dark pool is only used for $|x| > \frac{\Gamma}{2\Lambda}(T-t)$, and we expect the value function to be given by

$$w(t,x) = \begin{cases} C(t;0)x^2 = \frac{\Lambda}{T-t}x^2 & \text{if } |x| \le \frac{\Gamma}{2\Lambda}(T-t) \\ C_1(t,T)x^2 + C_2(t,T)|x| + C_3(t,T) & \text{if } |x| > \frac{\Gamma}{2\Lambda}(T-t). \end{cases} \tag{3.70}$$

As in Lemma 3.3.9 we obtain that

$$\frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s)^2 + \frac{\partial C_2}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_3}{\partial s}(t,s) = 0,$$

$$2\frac{\partial C_1}{\partial s}(t,s)\bar{X}(t,s) + \frac{\partial C_2}{\partial s}(t,s) = 0.$$

Thus, by Equation (3.69),

$$\frac{\partial}{\partial s}\big(C_1(t,s)\bar{X}(t,s)^2 + C_2(t,s)\bar{X}(t,s) + C_3(t,s)\big) = 0,$$

$$\frac{\partial}{\partial s}\big(2C_1(t,s)\bar{X}(t,s) + C_2(t,s)\big) = 0,$$

in particular

$$C_1(t,T)\Big(\frac{\Gamma}{2\Lambda}(T-t)\Big)^2 + C_2(t,T)\frac{\Gamma}{2\Lambda}(T-t) + C_3(t,T) = C(t;0)\Big(\frac{\Gamma}{2\Lambda}(T-t)\Big)^2,$$

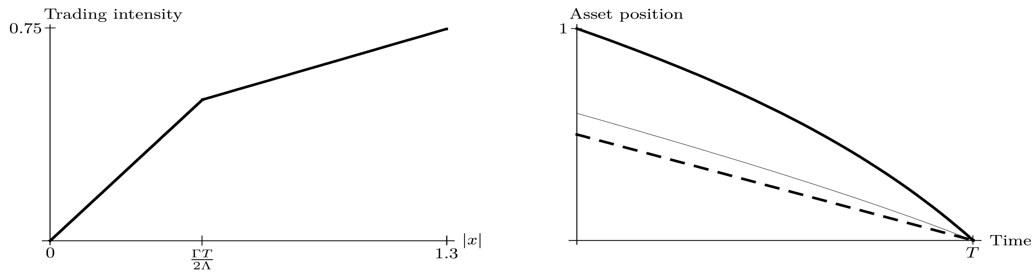


Figure 3.1.: The left picture shows the optimal trading intensity ξ^* at time $t = 0$ dependent on the initial asset position $x \in [0, 1.3]$. It is well visible that $\xi^*(0, \cdot)$ is not differentiable at $x = \frac{\Gamma T}{2\Lambda} = 0.5$. The right picture shows the optimal trading trajectories (provided there is no dark pool execution) for initial asset positions $x = 1$ (thick solid line) and $x = 0.6$ (thin solid line). The dashed line is the boundary $\frac{\Gamma(T-t)}{2\Lambda}$. Even for initial asset positions close to this boundary, the boundary is not crossed during the entire trading horizon unless the dark pool order is executed. After execution in the dark pool, the investor follows the dashed line for the remainder of the trading horizon. $T = \Lambda = \Gamma = 1$, $\theta = 2$.

$$2C_1(t,T)\frac{\Gamma}{2\Lambda}(T-t) + C_2(t,T) = 2C(t;0)\frac{\Gamma}{2\Lambda}(T-t).$$

We can deduce that

$$w(t,\cdot) \quad \text{and} \quad \frac{\partial w}{\partial x}(t,\cdot)$$

are continuous. However (cf. the proof of Theorem 3.3.10)

$$\frac{\partial^2 w}{\partial x^2}(t,x) = \begin{cases} 2C(t;0) & \text{if } |x| < \frac{\Gamma}{2\Lambda}(T-t) \\ 2C_1(t,T) < 2C(t;0) & \text{if } |x| > \frac{\Gamma}{2\Lambda}(T-t). \end{cases}$$

Defining the candidate optimal strategy as before by

$$\xi^* = \xi^*(t,x) := \frac{1}{2\Lambda}\frac{\partial w}{\partial x}(t,x) = \begin{cases} \frac{2C_1(t,T)x + \operatorname{sgn}(x)C_2(t,T)}{2\Lambda} & \text{if } |x| > \frac{\Gamma}{2\Lambda}(T-t) \\ \frac{x}{T-t} & \text{if } |x| \le \frac{\Gamma}{2\Lambda}(T-t), \end{cases} \tag{3.71}$$

$$\eta^* = \eta^*(t,x) := \begin{cases} \operatorname{sgn}(x)\Big(|x| - \frac{\Gamma}{2\Lambda}(T-t)\Big) & \text{if } |x| > \frac{\Gamma}{2\Lambda}(T-t) \\ 0 & \text{if } |x| \le \frac{\Gamma}{2\Lambda}(T-t), \end{cases} \tag{3.72}$$

we obtain that $\xi^*(t,\cdot)$ is continuous but *not* differentiable at

$$|x| = \frac{\Gamma}{2\Lambda}(T-t).$$

All steps of the proof of Theorem 3.4.4 can be replicated for the case $\alpha\Sigma = 0$ in a straightforward manner. The solution of the Optimization Problem ($\overline{\text{OPT}}$) is given by Equations (3.71) and (3.72). The value function is given by w as in Equation (3.70).

We want to conclude this section by illustrating the following two features of the optimal trading strategy in Figure 3.1. First, it is well visible that the optimal trading intensity $\xi^*(t,\cdot)$ is not differentiable at $x = \frac{\Gamma T}{2\Lambda}$. Secondly, if the initial asset position x at time t is larger than the boundary $\frac{\Gamma}{2\Lambda}T$, the optimal trading trajectory $X^*(s)$ never crosses the boundary $\frac{\Gamma}{2\Lambda}(T-s)$, provided there is no dark pool execution.

Let us briefly comment on this structure. For risk-neutral optimal liquidation without dark pool, the optimal trading intensity is constant (cf. the thin solid lines in Figure 1.3). The boundary itself is linear, and it is the trading trajectory of the optimal strategy (without dark pool) with initial position *on* the boundary. If the initial position is above the boundary, the usage of the dark pool *slows down* the optimal trading intensity, i.e., if the dark pool is not executed, the trading trajectory is concave and thus never crosses the boundary.

Conclusions

This thesis establishes a tractable mathematical benchmark model which captures the key properties of optimal order execution in dark pools.

The first part develops a discrete-time model which describes the stylized facts of both the dark pool and the traditional exchange and allows us to derive qualitative and quantitative characteristics of the optimal liquidation strategy and the value function of the respective optimization problems. In the second part we develop the mathematical tools for solving the stochastic control problems arising from this model in the continuous-time setting. The two parts are closely connected: the continuous-time optimization problems in Chapters 2 and 3 arise from discrete-time optimal control problems of Chapter 1. The structure of the solutions of these problems is analog to the structure of the solutions of the corresponding discrete-time models; especially in Chapter 3, we use the insight about the structure of the discrete-time solution in order to derive the continuous-time value function via heuristic considerations. Additionally, the solutions of the discrete-time optimization problems converge to the solutions of the continuous-time optimization problems under suitable assumptions.

The thesis contributes both to the economical and to the mathematical literature. We have outlined the economical contributions above: our market model may serve as the starting point for further research of optimal order execution in dark pools; these venues attain increasing importance and should not be ignored in this context. Mathematically, the solutions of the continuous-time optimization problems in Chapters 2 and 3 contribute to the literature on stochastic control:

- In Chapter 2 we solve the multidimensional linear-quadratic control problem with singular value function at terminal time T in a non-standard way. We bound positive definite solutions of a certain matrix differential equation by using a comparison result for Riccati matrix differential equations via an appropriate matrix inequality. We then use an approximation argument to show that the value function of the optimization problem is a quadratic form for a matrix-valued function which is the "principal solution" of the considered matrix differential equation.

- In Chapter 3 we solve the non-linear-quadratic one-dimensional control problem in *closed form* via heuristic considerations and a non-standard "interpolation procedure".

The thesis is by no means an exhausting treatment of the research topic. Both on the economical and on the mathematical side, there are numerous possibilities to generalize and/or modify the results obtained. We want conclude this thesis by mentioning a few problems which we think are interesting future research endeavors.

(i) In the general discrete-time optimization problem of Section 1.2 we required the additional assumptions that the probability space is finite and that trade execution in the dark pool is independent for the n assets (Assumption 1.2.1). While we believe that a finite probability space is not too restrictive, we think that it is desirable (and possible) to replace the second assumption by a generalization of the symmetry assumption we made in Section 1.3 (Assumption 1.3.1 (iv)). Note that removing the independence assumption without substitution, causes the convexity argument in the proof of Proposition 1.2.4 to break down (cf. Remark 1.3.2).

(ii) In Section 1.5 we analyze the effect of a different transaction price for the dark pool in a linear price impact model. If we assume that trades in the dark pool are settled at the price P prevailing at the primary venue, market manipulation might become possible and optimal strategies might not exist. We showed that this undesirable property disappears in the model specification of Section 1.5 if the dark pool liquidity is bounded and adverse selection is sufficiently strong. The conditions we find are rather strong and it is an interesting question how these can be weakened.

More broadly, our model is flexible enough to allow for a more general form of price impact at the primary exchange, e.g., in the form of temporary *and permanent* impact. If the transaction price in the dark pool comprises *partly* the price impact generated in the primary venue, e.g., the permanent impact but not the temporary impact, similar effects as in Section 1.5 might occur. It is thus also interesting to analyze modified settings in the context of market manipulation in dark pools.

(iii) In the discrete-time model the criterion for the costs of a trading strategy was the implementation shortfall. In the model specifications of Sections 1.3 and 1.4 the fundamental asset price $\tilde{P}$ is a martingale and the cost functional is independent of its development. In the corresponding continuous-time optimization problems we therefore do not include $\tilde{P}$ into the cost functionals. If we analyze the minimization of the implementation shortfall instead, this could lead to technical and mathematical difficulties even if $\tilde{P}$ is a martingale which we leave for future research.

(iv) In the continuous-time model of Chapter 2, the price impact matrix Λ and the covariance matrix Σ are constant in time. For the case $n = 1$, the argumentation of Chapter 2 can be replicated for time-dependent functions $\Sigma(t)$ and $\Lambda(t)$ (if these are, e.g., continuous). Note that closed form solutions for the value function and the optimal strategy are not possible anymore as the value function is no longer the solution of a scalar Riccati equation with *constant* coefficients.

For general n, we can still replace Σ by a time-dependent matrix-valued function $\Sigma(t)$ without changing the argumentation essentially. It is an open question how the results can be transfered to a time-dependent function $\Lambda(t)$ instead of Λ.

(v) In Section 2.6 we showed that the value function and the optimal strategy of the *single asset* discrete-time optimization problem of Section 1.3 converge to the corresponding objects of the continuous-time optimization problem of Chapter 2 if the number of trading times tends to infinity and the step size is constant (provided the parameters are scaled appropriately). In these cases we have closed form solutions both in the discrete and in the continuous-time setting. We are not yet able to give a rigorous proof for the respective *multi asset* result (Conjecture 2.6.2).

We want to remark here that a convergence result for the models with adverse selection is possible as the discrete and the continuous-time optimization problems were solved in closed form. Because of the rather complex formulae, we omitted the tedious execution.

(vi) We only treated adverse selection for *single asset* liquidation. Given the complexity of the heuristic derivation of the value function of the optimization problems both in the discrete-time and in the continuous-time setting, we believe that it is difficult to solve this problem. Nevertheless, we believe that it is one of the most interesting research questions raised by this thesis.

A. Appendix

A.1. Recursions for Section 1.4.4

The functions

$$A_1, A_2, B_1, B_2, C_1, C_2, C_3 : \{(t,s) \in \mathbb{R}^2 \,|\, t, s \in \{t_0, \dots t_N\}, t \leq s\} \longrightarrow \mathbb{R}$$

are given recursively by

$$C_1(t_i, t_i) = C(t_i, 0), \quad C_2(t_i, t_i) = C_3(t_i, t_i) = 0, \tag{A.1}$$
$$A_1(t_i, t_i) = A(t_i, 0), \quad A_2(t_i, t_i) = 0, \tag{A.2}$$
$$B_1(t_i, t_i) = B_2(t_i, t_i) = 0$$

for $i = 0, \dots, N$ and

$$C_1(t_i, t_j) = \frac{\alpha\Sigma\Lambda + (1-p)C_1(t_{i+1}, t_j)(\Lambda + \alpha\Sigma)}{C_1(t_{i+1}, t_j)(1-p) + \Lambda}, \tag{A.3}$$
$$C_2(t_i, t_j) = \frac{((1-p)C_2(t_{i+1}, t_j) + \Gamma p)\Lambda}{(1-p)C_1(t_{i+1}, t_j) + \Lambda}, \tag{A.4}$$
$$C_3(t_i, t_j) = (1-p)C_3(t_{i+1}, t_j) - \frac{\Gamma^2 p}{4C(t_{i+1}, 0)} - \frac{\left((1-p)C_2(t_{i+1}, t_j) + \Gamma p\right)^2}{4((1-p)C_1(t_{i+1}, t_j) + \Lambda)}, \tag{A.5}$$
$$A_1(t_i, t_j) = \frac{(1-p)C_1(t_{i+1}, t_j)}{(1-p)C_1(t_{i+1}, t_j) + \Lambda}, \tag{A.6}$$
$$A_2(t_i, t_j) = \frac{(1-p)C_2(t_{i+1}, t_j) + \Gamma p}{2((1-p)C_1(t_{i+1}, t_j) + \Lambda)}, \tag{A.7}$$
$$B_1(t_i, t_j) = \frac{\Lambda}{(1-p)C_1(t_{i+1}, t_j) + \Lambda}, \tag{A.8}$$
$$B_2(t_i, t_j) = -\frac{(1-p)C(t_{i+1},0)C_2(t_{i+1},t_j) + \Gamma\left((1-p)C_1(t_{i+1},t_j) + C(t_{i+1},0)p + \Lambda\right)}{2C(t_{i+1},0)(C_1(t_{i+1},t_j)(1-p) + \Lambda)} \tag{A.9}$$

for $i = 0, \dots, N-1$, $j = i+1, \dots, N$.

The function

$$\bar{X} : \{(t,s) \in \mathbb{R}^2 \,|\, t, s \in \{t_0, \dots t_N\}, t < s\} \longrightarrow \mathbb{R},$$

is given recursively by

$$\bar{X}(t_i, t_{i+1}) = \frac{\Gamma(C(t_{i+1}, 0) + \Lambda)}{2C(t_{i+1}, 0)\Lambda} \tag{A.10}$$

for $i = 0, \dots, N-1$ and

$$\bar{X}(t_i, t_j) = \frac{C_2(t_{i+1}, t_{j-1})(1-p) + \Gamma p + 2(C_1(t_{i+1}, t_{j-1})(1-p) + \Lambda)\bar{X}(t_{i+1}, t_j)}{2\Lambda} \quad \text{(A.11)}$$

for $i = 0, \dots, N-2$ and $j = i+2, ..., N$.

A.2. Closed form solutions for Section 1.4.6

Let $i = 0, \dots, N-1$, $j = i+1, \dots, N-1$. By abuse of notation, we define

$$\bar{X}(t_{i+1}, t_{i+1}) := \frac{\Gamma}{2C(t_{i+2}, 0)} \neq 0.$$

Then

$$A_1(t_i, t_j) = 1 - \frac{\bar{X}(t_{i+1}, t_{j+1}) - \bar{X}(t_{i+1}, t_j)}{\bar{X}(t_i, t_{j+1}) - \bar{X}(t_i, t_j)}, \quad \text{(A.12)}$$

$$A_2(t_i, t_j) = \frac{\bar{X}(t_{i+1}, t_{j+1})\bar{X}_{i,j} - \bar{X}(t_{i+1}, t_j)\bar{X}(t_i, t_{j+1})}{\bar{X}(t_i, t_{j+1}) - \bar{X}_{i,j}}, \quad \text{(A.13)}$$

$$B_1(t_i, t_j) = 1 - A_1(t_i, t_j), \quad \text{(A.14)}$$

$$B_2(t_i, t_j) = -A_2(t_i, t_j) - \frac{\Gamma}{2C(t_{i+1}, 0)}, \quad \text{(A.15)}$$

$$C_1(t_i, t_j) = \frac{\Lambda}{1-p}\left(\frac{\bar{X}(t_{i-1}, t_{j+1}) - \bar{X}(t_{i-1}, t_j)}{\bar{X}(t_i, t_{j+1}) - \bar{X}(t_i, t_j)} - 1\right), \quad \text{(A.16)}$$

$$C_2(t_i, t_j) = \frac{\bar{X}(t_i, t_{j+1})(2\Lambda\bar{X}(t_{i-1}, t_j) - \Gamma p) - \bar{X}(t_i, t_j)(2\Lambda\bar{X}(t_{i-1}, t_{j+1}) - \Gamma p)}{(1-p)(\bar{X}(t_i, t_{j+1}) - \bar{X}(t_i, t_j))}, \quad \text{(A.17)}$$

where $\bar{X}$ is given as in Equation (1.76).

Furthermore, for $i = 0, \dots, N-1$,

$$A_1(t_i, t_N) = A(t_i, p), \quad \text{(A.18)}$$

$$A_2(t_i, t_N) = B(t_i, p)\bar{X}(t_i, t_N) - \bar{X}(t_{i+1}, t_N), \quad \text{(A.19)}$$

$$B_1(t_i, t_N) = B(t_i, p), \quad \text{(A.20)}$$

$$B_2(t_i, t_N) = -A_2(t_i, t_j) - \frac{\Gamma}{2C(t_{i+1}, 0)}, \quad \text{(A.21)}$$

$$C_1(t_i, t_N) = C(t_i, p), \quad \text{(A.22)}$$

$$C_2(t_i, t_N) = \frac{A_2(t_{i-1}, t_N)(C(t_i, p)(1-p) + \Lambda)}{1-p}. \quad \text{(A.23)}$$

Finally, for $i, \dots, N-1$, $j = i+1, \dots, N$, we have

$$C_3(t_i,t_j) = \sum_{k=i}^{j-1} (1-p)^{k-i} D(t_{k+1},t_j), \quad \text{where} \tag{A.24}$$

$$D(t_i,t_j) := -\frac{\Gamma^2 p}{4C_i(0)} - \frac{\Big(C_2(t_i,t_j)(1-p)+\Gamma p\Big)^2}{4(C_1(t_i,t_j)(1-p)+\Lambda)}.$$

A.3. Detailed proofs for selected formulae of Chapter 3

We give proofs for selected formulae which were not executed in the main body of the thesis. All calculations have been carried out by hand and afterwards been verified with *Mathematica*.

Proof of Equation (3.30). Note first that

$$\int_s^t \Big(\frac{C_1(u,s)}{\Lambda} + \theta\Big) du = \Big[-\log\big(\sinh\big(\frac{\tilde\theta}{2}(s-u)+\kappa(s)\big)\big) + \frac{\theta}{2}u\Big]_s^t$$

$$= -\log\big(\sinh\big(\frac{\tilde\theta}{2}(s-t)+\kappa(s)\big)\big) + \log\big(\sinh(\kappa(s))\big) + \frac{\theta}{2}(t-s).$$

Therefore (note that $\tilde\theta^2 - \theta^2 = \frac{4\alpha\Sigma}{\Lambda}$),

$$\begin{aligned}
C_2(t,s) &= -\theta\Gamma \frac{\sinh(\kappa(s))\exp\big(\frac{\theta}{2}(t-s)\big)}{\sinh\big(\frac{\tilde\theta}{2}(s-t)+\kappa(s)\big)} \int_s^t \frac{\sinh\big(\frac{\tilde\theta}{2}(s-u)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(s-u)\big)}{\sinh(\kappa(s))} du \\
&= -\theta\Gamma \frac{\exp\big(\frac{\theta}{2}(t-s)\big)}{\sinh\big(\frac{\tilde\theta}{2}(s-t)+\kappa(s)\big)} \\
&\qquad \Big[4\frac{\exp\big(\frac{\theta}{2}(s-u)\big)\big(\frac{\tilde\theta}{2}\cosh\big(\frac{\tilde\theta}{2}(s-u)+\kappa(s)\big) - \frac{\theta}{2}\sinh\big(\frac{\tilde\theta}{2}(s-u)+\kappa(s)\big)\big)}{\tilde\theta^2-\theta^2}\Big]_t^s \\
&= \frac{\theta\Gamma\Lambda}{2\alpha\Sigma}\Big(- \frac{\exp\big(\frac{\theta}{2}(t-s)\big)\big(\tilde\theta\cosh(\kappa(s)) - \theta\sinh(\kappa(s))\big)}{\sinh\big(\frac{\tilde\theta}{2}(s-t)+\kappa(s)\big)} \\
&\qquad + \tilde\theta\coth\big(\frac{\tilde\theta}{2}(s-t)+\kappa(s)\big) - \theta\Big).
\end{aligned} \tag{A.25}$$

We have

$$\sinh(\kappa(s)) = \sinh(\operatorname{arcoth}(\mu(s))) = \frac{1}{\sqrt{\mu(s)^2-1}}, \tag{A.26}$$

$$\cosh(\kappa(s)) = \sinh(\operatorname{arcoth}(\mu(s))) = \frac{\mu(s)}{\sqrt{\mu(s)^2-1}} \tag{A.27}$$

and therefore

$$\tilde{\theta}\cosh(\kappa(s)) - \theta\sinh(\kappa(s)) = \frac{2C(s;0)}{\Lambda\sqrt{\mu(s)^2-1}}.$$

Plugging this into Equation (A.25), we obtain

$$\begin{aligned}C_2(t,s) = \frac{\theta\Gamma\Lambda}{2\alpha\Sigma}\Big(&\tilde{\theta}\coth\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\\ &- \frac{2C(s;0)}{\Lambda\sqrt{\mu(s)^2-1}}\frac{\exp\big(\frac{\theta}{2}(t-s)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)} - \theta\Big).\end{aligned} \tag{A.28}$$

We use the addition formulae for hyperbolic functions and Equations (A.26) and (A.27) again and conclude as follows:

$$\begin{aligned}C_2(t,s) &= \frac{\theta\Gamma\Lambda}{2\alpha\Sigma}\Big(\tilde{\theta}\frac{1+\coth\big(\frac{\tilde{\theta}}{2}(s-t)\big)\coth(\kappa(s))}{\coth\big(\frac{\tilde{\theta}}{2}(s-t)\big)+\coth(\kappa(s))}\\ &\quad - \frac{2C(s;0)}{\Lambda\sqrt{\mu(s)^2-1}}\frac{\exp\big(\frac{\theta}{2}(t-s)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\cosh(\kappa(s))+\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\sinh(\kappa(s))} - \theta\Big)\\ &= \frac{\theta\Gamma}{2\alpha\Sigma}\Big(\Lambda\tilde{\theta}\frac{\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)+\mu(s)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)}{\mu(s)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)+\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)}\\ &\quad - \frac{2C(s;0)\exp\big(\frac{\theta}{2}(t-s)\big)}{\mu(s)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)+\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)} - \Lambda\theta\Big).\end{aligned}$$

□

Proof of Equation (3.32). Similarly as before, we obtain

$$\int_t^s \Big(\frac{C_1(u,s)}{\Lambda}\Big)du = \log\big(\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\big) - \log\big(\sinh(\kappa(s))\big) + \frac{\theta}{2}(t-s).$$

Furthermore, by partial integration, we have

$$\begin{aligned}\int_s^t \frac{\theta\exp\big(\frac{\theta}{2}(s-u)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)}du &= \theta\int_s^t \exp\big(\frac{\theta}{2}(s-u)\big)\frac{1}{\sinh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)}du\\ &= -2\Big[\frac{\exp\big(\frac{\theta}{2}(s-u)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)}\Big]_s^t\\ &\quad + \tilde{\theta}\int_s^t \frac{\exp\big(\frac{\theta}{2}(s-u)\big)\cosh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)}du.\end{aligned} \tag{A.29}$$

Thus, by Equation (A.28),

$$\begin{aligned}
\bar{X}(t,s) &= \frac{\Gamma}{2C(s;0)} \frac{\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big)}{\sinh(\kappa(s))} \\
&\quad - \frac{\theta\Gamma}{4\alpha\Sigma} \frac{\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big)}{\sinh(\kappa(s))} \\
&\quad \int_s^t \Big(\frac{\sinh(\kappa(s))\exp\big(\frac{\theta}{2}(s-u)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)} \Big(\tilde{\theta}\coth\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big) \\
&\qquad - \frac{2C(s;0)}{\Lambda\sqrt{\mu(s)^2-1}} \frac{\exp\big(\frac{\theta}{2}(u-s)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(u-t)+\kappa(s)\big)} - \theta\Big)\Big) du \\
&= \frac{\Gamma\sqrt{\mu(s)^2-1}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big)}{2C(s;0)} \\
&\quad - \frac{\theta\Gamma}{4\alpha\Sigma}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big) \\
&\quad \int_s^t \Big(\frac{\tilde{\theta}\cosh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(s-u)\big)}{\sinh^2\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)} - \frac{\theta\exp\big(\frac{\theta}{2}(s-u)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)} \\
&\qquad - \frac{2C(s;0)}{\Lambda\sqrt{\mu(s)^2-1}\sinh^2\big(\frac{\tilde{\theta}}{2}(u-t)+\kappa(s)\big)}\Big) du \\
&\overset{(A.29)}{=} \frac{\Gamma\sqrt{\mu(s)^2-1}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big)}{2C(s;0)} \\
&\quad - \frac{\theta\Gamma}{4\alpha\Sigma}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big) \\
&\quad \Big(2\Big[\frac{\exp\big(\frac{\theta}{2}(s-u)\big)}{\sinh\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)}\Big]_s^t - \frac{4C(s;0)}{\Lambda\tilde{\theta}\sqrt{\mu(s)^2-1}}\Big[\coth\big(\frac{\tilde{\theta}}{2}(s-u)+\kappa(s)\big)\Big]_s^t\Big) \\
&= \frac{\Gamma\sqrt{\mu(s)^2-1}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big)}{2C(s;0)} - \frac{\theta\Gamma}{2\alpha\Sigma} \\
&\quad - \frac{\theta\Gamma C(s;0)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big)}{\alpha\Sigma\Lambda\tilde{\theta}\sqrt{\mu(s)^2-1}} \\
&\quad - \frac{\big(2\alpha\Sigma\Lambda\theta\Gamma - \Lambda\theta^2\Gamma C(s;0)\big)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)\exp\big(\frac{\theta}{2}(t-s)\big)}{\alpha\Sigma\Lambda^2\tilde{\theta}^2\sqrt{\mu(s)^2-1}},
\end{aligned} \tag{A.30}$$

where we used the fact that

$$\mu(s)^2 - 1 = 4\frac{C(s;0)^2 + \Lambda\theta C(s;0) - \alpha\Sigma\Lambda}{\Lambda^2\tilde{\theta}^2} \tag{A.31}$$

in Equation (A.30). Using the addition formulae for hyperbolic functions and Equations (A.26), (A.27) and (A.31), we obtain Equation (3.32). □

Proof of Equation (3.39). Note first that $C(\cdot;0)$ and thus μ is differentiable in $(0,T)$ and that

$$C'(s;0) = \frac{C(s;0)^2}{\Lambda} - \alpha\Sigma. \tag{A.32}$$

We compute

$$\begin{aligned}
\frac{\partial \bar{X}}{\partial s}(t,s) = &\frac{\partial}{\partial s}\Big(\frac{\Gamma\mu(s)}{2C(s;0)}\Big)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\theta}{2}(t-s)\big)\\
&+\frac{\partial}{\partial s}\Big(\frac{\Gamma}{2C(s;0)}\Big)\exp\big(\frac{\theta}{2}(t-s)\big)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\\
&+\exp\big(\frac{\theta}{2}(t-s)\big)\Big(\frac{\Gamma\mu(s)}{2C(s;0)}+\frac{\theta^2\Gamma}{2\tilde{\theta}\alpha\Sigma}\Big)\Big(\frac{\tilde{\theta}}{2}\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)-\frac{\theta}{2}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\Big)\\
&+\exp\big(\frac{\theta}{2}(t-s)\big)\Big(\frac{\Gamma}{2C(s;0)}+\frac{\theta\Gamma}{2\alpha\Sigma}\Big)\Big(\frac{\tilde{\theta}}{2}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)-\frac{\theta}{2}\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\Big)\\
=:{}& a(s)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\theta}{2}(t-s)\big)+b(s)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\theta}{2}(t-s)\big).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
\frac{\mu(s)}{C(s;0)} &= \frac{2}{\Lambda\tilde{\theta}}+\frac{\theta}{\tilde{\theta}C(0;s)},\\
\frac{\partial}{\partial s}\Big(\frac{\mu(s)}{C(s;0)}\Big) &= -\frac{\theta C'(s;0)}{\tilde{\theta}C(s;0)^2} = -\frac{\theta}{\Lambda\tilde{\theta}}+\frac{\theta\alpha\Sigma}{\tilde{\theta}C(0;s)^2},\\
\frac{\partial}{\partial s}\Big(\frac{1}{C(s;0)}\Big) &= -\frac{\theta C'(s;0)}{\tilde{\theta}C(s;0)^2} = -\frac{1}{\Lambda}+\frac{\alpha\Sigma}{C(s;0)^2}.
\end{aligned}$$

Therefore (note again that $\tilde{\theta}^2-\theta^2 = \frac{4\alpha\Sigma}{\Lambda}$),

$$\begin{aligned}
a(s) &= -\frac{\theta\Gamma}{2\Lambda\tilde{\theta}}+\frac{\theta\alpha\Sigma\Gamma}{2\tilde{\theta}C(s;0)^2}-\frac{\theta\Gamma}{2\Lambda\tilde{\theta}}-\frac{\theta^2\Gamma}{4\tilde{\theta}C(s;0)}-\frac{\theta^3\Gamma}{4\tilde{\theta}\alpha\Sigma}+\frac{\tilde{\theta}\Gamma}{4C(s;0)}+\frac{\theta\tilde{\theta}\Gamma}{4\alpha\Sigma}\\
&= -\frac{\theta\Gamma}{\Lambda\tilde{\theta}}+\frac{\theta\alpha\Sigma\Lambda\Gamma}{2\Lambda\tilde{\theta}C(s;0)^2}\underbrace{-\frac{\Lambda\theta^2\Gamma}{4\Lambda\tilde{\theta}C(s;0)}+\frac{\Lambda\tilde{\theta}^2\Gamma}{4\Lambda\tilde{\theta}C(s;0)}}_{=\frac{\alpha\Sigma\Gamma}{\Lambda\tilde{\theta}C(s;0)}}\underbrace{-\frac{\theta^3\Gamma}{4\tilde{\theta}\alpha\Sigma}+\frac{\theta\tilde{\theta}^2\Gamma}{4\tilde{\theta}\alpha\Sigma}}_{=\frac{\theta\Gamma}{\Lambda\tilde{\theta}}}\\
&= \frac{\alpha\Sigma\Gamma}{2C(s;0)^2}\Big(\frac{\Lambda\theta+2C(s;0)}{\Lambda\tilde{\theta}}\Big) = \frac{\alpha\Sigma\Gamma}{2C(s;0)^2}\mu(s),\\
b(s) &= \frac{\Gamma}{2\Lambda}+\frac{\theta^2\Gamma}{4\alpha\Sigma}-\frac{\Gamma}{2\Lambda}+\frac{\alpha\Sigma\Gamma}{2C(s;0)^2}-\frac{\theta\Gamma}{4C(s;0)}-\frac{\theta^2\Gamma}{4\alpha\Sigma}+\frac{\theta\Gamma}{4C(s;0)}\\
&= \frac{\alpha\Sigma\Gamma}{2C(s;0)^2}
\end{aligned}$$

and Equation (3.39) follows directly. □

Proof of Equation (3.44). Note first that

$$\mu'(s) = \frac{2C(s;0)^2 - 2\alpha\Sigma\Lambda}{\Lambda^2\tilde{\theta}}.$$

Furthermore by Equation (A.31),

$$\begin{aligned}
\kappa'(s) + \frac{\tilde{\theta}}{2} &= \frac{\mu'(s)}{1-\mu(s)^2} + \frac{\tilde{\theta}}{2} \\
&= \frac{4C(s;0)^2 - 4\alpha\Sigma\Lambda + \Lambda^2\tilde{\theta}^2(1-\mu(s)^2)}{2\Lambda^2\tilde{\theta}(1-\mu(s)^2)} \\
&= \frac{-4\theta C(s;0)}{2\Lambda\tilde{\theta}(1-\mu(s)^2)}.
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{\partial C_1}{\partial s}(t,s) &= -\frac{\Lambda\tilde{\theta}}{2}\Big(\kappa'(s) + \frac{\tilde{\theta}}{2}\Big)\frac{1}{\sinh^2\big(\frac{\tilde{\theta}}{2}(s-t) + \kappa(s)\big)} \\
&= \frac{\theta C(s;0)}{(1-\mu(s)^2)\sinh^2\big(\frac{\tilde{\theta}}{2}(s-t)+\kappa(s)\big)} < 0
\end{aligned}$$

as $\mu(s) > 1$. Using the addition formulae for hyperbolic functions, Equation (3.44) follows. □

Proof of Equation (3.45). Note first that

$$\mu'(s) = \frac{2C'(s;0)}{\tilde{\theta}\Lambda}.$$

Let now

$$\begin{aligned}
f(s) &:= \mu(s)\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big) + \cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big), \\
g(s) &:= \Lambda\tilde{\theta}\sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big) + \Lambda\tilde{\theta}\mu(s)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big) - 2C(s;0)\exp\big(\frac{\theta}{2}(t-s)\big).
\end{aligned}$$

Using Equation (A.32), we have

$$\theta\mu(s)C(s;0) - 2\mu(s)C(s;0) + \frac{4C(s;0)C'(s;0)}{\Lambda\tilde{\theta}} + \tilde{\theta}C(s;0) = \frac{2\tilde{\theta}\alpha\Sigma + (\theta^2+\tilde{\theta}^2)C(s;0)}{\tilde{\theta}}$$

$$\theta\mu(s)C(s;0)\tilde{\theta}\mu(s)C(s;0) - 2C'(s;0) + \theta C(s;0) = 2\theta C(s;0) + 2\alpha\Sigma.$$

Thus, we are able to compute (note that $\sinh^2(x) - \cosh(x) = -1$)

$$\begin{aligned}
(fg' - gf')(s) = &\Big(\frac{\Lambda\tilde{\theta}^2(\mu(s)^2 - 1)}{2} - 2C'(s;0)\Big) \cdot \sinh^2\big(\frac{\tilde{\theta}}{2}(s-t)\big) \\
&+ \Big(\frac{\Lambda\tilde{\theta}^2(1 - \mu(s)^2)}{2} + 2C'(s;0)\Big) \cdot \cosh^2\big(\frac{\tilde{\theta}}{2}(s-t)\big) \\
&+ \Big(\theta\mu(s)C(s;0) - 2\mu(s)C(s;0) + \frac{4C(s;0)C'(s;0)}{\Lambda\tilde{\theta}} + \tilde{\theta}C(s;0)\Big) \\
&\quad \sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\tilde{\theta}}{2}(t-s)\big) \\
&+ \Big(\theta\mu(s)C(s;0) - 2C'(s;0) + \theta C(s;0)\Big) \\
&\quad \cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\tilde{\theta}}{2}(t-s)\big) \\
= &-2\theta C(s;0) + \frac{2\tilde{\theta}\alpha\Sigma + (\theta^2 + \tilde{\theta}^2)C(s;0)}{\tilde{\theta}} \sinh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\tilde{\theta}}{2}(t-s)\big) \\
&+ \Big(2\theta C(s;0) + 2\alpha\Sigma\Big)\cosh\big(\frac{\tilde{\theta}}{2}(s-t)\big)\exp\big(\frac{\tilde{\theta}}{2}(t-s)\big).
\end{aligned}$$

Equation (3.45) follows directly. □

Bibliography

P. Abramowitz. Tool of the trade. *Institutional Investor's Alpha Magazine*, 6:41–44, 2006.

A. Alfonsi and A. Schied. Optimal trade execution and absence of price manipulations in limit order book models. *SIAM J. Financial Math.*, 1:490–522, 2010.

A. Alfonsi, A. Schied, and A. Slynko. Order book resilience, price manipulation, and the positive portfolio problem. *Working Paper*, 2009.

A. Alfonsi, A. Fruth, and A. Schied. Optimal execution strategies in limit order books with general shape functions. *Quant. Finance*, 10(2):143–157, 2010.

R. Almgren. Optimal execution with nonlinear impact functions and trading-enhanced risk. *Appl. Math. Finance*, 10(1):1–18, 2003.

R. Almgren and N. Chriss. Optimal execution of portfolio transactions. *J. Risk*, 3(2): 5–39, 2001.

R. Almgren and J. Lorenz. Adaptive arrival price. *Algorithmic Trading III*, pages 59–66, 2007.

D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge University Press, Cambridge, 2004.

A. Ben-Israel and T. N. E. Greville. *Generalized inverses: Theory and applications.* Springer, New York, 2003.

D. S. Bernstein. *Matrix mathematics: Theory, facts, and formulas with application to linear systems theory.* Princeton University Press, Princeton, 2005.

D. Bertsimas and A. Lo. Optimal control of execution costs. *J. Financial Markets*, 1 (1):1–50, 1998.

B. I. Carlin, M. Sousa Lobo, and S. Viswanathan. Episodic liquidity crises: Cooperative and predatory trading. *J. Finance*, 65(5):2235–2274, 2007.

C. Carrie. Illuminating the new dark influence on trading and U.S. market structure. *J. Trading*, 3(2):40–55, 2008.

J. Conrad, K. M. Johnson, and S. Wahal. Institutional trading and alternative trading systems. *J. Finan. Econ.*, 70(1):99–134, 2003.

W. A. Coppel. *Disconjugacy.* Lecture Notes in Mathematics, Vol. 220. Springer, Berlin, 1971.

R. Courant and D. Hilbert. *Methods of mathematical physics, Volume I.* Wiley, New York, 1953.

H. Degryse, M. Van Achter, and G. Wuyts. Dynamic order submission strategies with competition between a dealer market and a crossing network. *J. Finan. Econ.*, 91(3): 319–338, 2009.

K. Fong, A. Madhavan, and P. L. Swan. Upstairs, downstairs: Does the upstairs market hurt the downstairs? *Working paper*, 2004.

K. Ganchev, M. Kearns, Y. Nevmyvaka, and J. Wortman Vaughan. Censored exploration and the dark pool problem. *Commun. ACM*, 53(5):99–107, 2010.

J. Gatheral. No-dynamic-arbitrage and market impact. *Quant. Finance*, 10(7):749–759, 2010.

T. Hendershott and H. Mendelson. Crossing networks and dealer markets: Competition and performance. *Journal of Finance*, 55(5):2071–2115, 2000.

R. A. Horn and C. R. Johnson. *Matrix analysis.* Cambridge University Press, Cambridge, 1985.

R. A. Horn and C. R. Johnson. *Topics in matrix analysis.* Cambridge University Press, Cambridge, 1991.

M. Höschler. *Limit order book models and optimal trading strategies.* PhD thesis, Technische Universität Berlin, 2011. In preparation.

G. Huberman and W. Stanzl. Price manipulation and quasi-arbitrage. *Econometrica*, 72(4):1247–1275, 2004.

O. Kallenberg. *Foundations of modern probability.* Springer, New York, 2002.

R. Kissell and M. Glantz. *Optimal Trading Strategies: Quantitative Approaches for Managing Market Impact and Trading Risk.* AMACOM Inc., New York, 2003.

P. Kratz and T. Schöneborn. Optimal liquidation in dark pools. *Working paper*, 2010.

S. Laruelle, C.-A. Lehalle, and G. Pagès. Optimal split of orders across liquidity pools: A stochastic algorithm approach. *Working papaer*, 2009.

D. Leinweber. Algo vs. algo. *Institutional Investor's Alpha Magazine*, 2:44–51, 2007.

H. Mittal. Are you playing in a toxic dark pool? A guide to preventing information leakage. *J. Trading*, pages 20–33, 2008.

R. Næs and B. A. Ødegaard. Equity trading by institutional investors. To cross or not to cross? *J. Financial Markets*, 9(2):79–99, 2006.

R. Næs and J. A. Skjeltorp. Equity trading by institutional investors: Evidence on order submission strategies. *J. Banking Finance*, 27(9):1779–1817, 2003.

F. Naujokat and N. Westray. Curve following in limit order markets. *Working paper*, 2010.

A. Obizhaeva and J. Wang. Optimal trading strategy and supply/demand dynamics. *To appear in J. Financial Markets*, 2006.

B. Øksendal. *Stochastic differential equations: An introduction with applications.* Springer, Berlin, 2007.

B. Øksendal and Y. Hu. Partial information linear quadratic control for jump diffusions. *SIAM J. Control Optim.*, 47(4):1744–1761, 2008.

B. Øksendal and A. Sulem. *Applied stochastic control of jump diffusions.* Springer, Berlin, 2007.

A. F. Perold. The implementation shortfall: Paper versus reality. *J. Portfol. Manage.*, 14(2):4–9, 1988.

O. Perron. *Die Lehre von den Kettenbrüchen. Bd I. Elementare Kettenbrüche.* Teubner, Stuttgart, 1954.

W. T. Reid. *Riccati differential equations.* Academic Press, New York, 1972.

L. C. G. Rogers and S. Singh. The cost of illiquidity and its effects on hedging. *Math. Finance*, 20(4):597 – 615, 2010.

J. Schack. The orders of battle. *Institutional Investor*, 11:77–84, 2004.

A. Schied and T. Schöneborn. Risk aversion and the dynamics of optimal liquidation strategies in illiquid markets. *Finance Stoch.*, 13(2):181–204, 2009.

A. Schied, T. Schöneborn, and M. Tehranchi. Optimal basket liquidation for CARA investors is deterministic. *Appl. Math. Finance*, 17(6):471–489, 2010.

T. Schöneborn and A. Schied. Liquidation in the face of adversity: Stealth vs. sunshine trading. *Working paper*, 2009.

Index of Notation

n	number of assets
T	time-horizon
N	number of trading intervals
t_i	trading time
$(\Omega, \mathcal{F}, \mathbb{P})$	underlying probability space
ω, $\omega_l, \ldots$	elements of Ω
$\mathbb{F} = (\mathcal{F}_{t_i})_{i=0,\ldots,N}$	filtration reflecting the total information in the market
$\tilde{\mathbb{F}} = (\tilde{\mathcal{F}}_{t_i})_{i=0,\ldots,N}$	filtration reflecting the information of the investor
$\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$	filtration in continuous time
$\mathbb{E}$	expectation with respect to $\mathbb{P}$
$\mathbb{E}_{t,x}$	expectation conditional on portfolio position x at time t
λ	Lebesgue measure on $[0,T]$ or $[t,T]$
$\tilde{P}(t_i)$	fundamental asset price
$P(t_i)$	price at the primary exchange comprising the price impact of the investor
$\epsilon(t_i)$	increments of $\tilde{P}$
$\Sigma(t_i)$, Σ	covariance matrix
α	risk-aversion parameter
f_i	price impact function
Λ	price impact matrix
$D := \sqrt{\Lambda^{-1}}\Sigma\sqrt{\Lambda^{-1}}$	
$A > 0,\ A \geq 0$	A is positive respectively nonnegative definite
Γ	adverse selection parameter
L	liquidity bound
$a(t_i)$, $b(t_i)$	dark pool liquidity; supply respectively demand
p_l	probability of a specific "scenario" for multi asset liquidation
Z_l	diagonal matrix with 1's and 0's on the diagonal associated with scenario l
$\check{M} := \sum_l p_l Z_l M Z_l$	
$\hat{P} := \sum_l p_l Z_l$	
$M^\dagger$	Moore-Penrose inverse of M

p	probability of execution for single asset liquidation
$\kappa(p) := \operatorname{arcosh}\Big(\frac{\sqrt{1-p}}{2}\big(\frac{\alpha\Sigma}{\Lambda}+1+\frac{1}{1-p}\big)\Big)$	
$\pi = (\pi_1, \dots, \pi_n)$	n-dimensional Poisson process governing dark pool execution
θ_i	intensity of the Poisson process π_i
$\theta := \sum_i \theta_i$	
$\tilde{\theta} := \sqrt{\theta^2 + \frac{4\alpha\Sigma}{\Lambda}}$	
e_i	i^{th} unit vector
τ_i	i^{th} jump time of the Poisson process π in $[0,T]$ or $[t,T]$
$\mathbb{A}(t_i, X(t_i))$	set of admissible liquidation strategies in discrete time
$\tilde{\mathbb{A}}(t,x)$	set of admissible strategies in continuous time
$\mathbb{A}(t,x)$	set of admissible liquidation strategies in continuous time
$\mathcal{R}(\cdot)$	implementation shortfall
$J(\cdot)$, $\bar{J}(\cdot)$, $\tilde{J}(\cdot)$	cost functionals
$f(\cdot)$, $\bar{f}(\cdot)$	running costs in continuous time
$v(t_i, x(t_0), \dots, x(t_{i-1}), X(t_i))$, $v(t_i, X(t_i))$, $\bar{v}(t_i, X(t_i))$	value functions in discrete time
$v(t,x)$, $\tilde{v}(t,x)$, $\bar{v}(t,x)$	value functions in continuous time
$w(t,x)$	candidate value function in continuous time
$(x(t_i))_{i=0,\dots,N}$	trading strategy in the primary venue
$(y(t_i))_{i=0,\dots,N}$	trading strategy in the dark pool
$(z(t_i))_{i=0,\dots,N}$	executed dark pool orders associated with $(y(t_i))_{i=0,\dots,N}$
$(x^*(t_i))_{i=0,\dots,N}$	optimal trading strategy in the primary venue
$(y^*(t_i))_{i=0,\dots,N}$	optimal trading strategy in the dark pool
$(z^*(t_i))_{i=0,\dots,N}$	executed dark pool orders associated with $(y^*(t_i))_{i=0,\dots,N}$
$X(t_i)$	portfolio position at time t_i associated with the trading strategy $((x(t_i), y(t_i)))_{i=0,\dots,N}$
$X^{\text{ne}}(t_i)$	optimal trading trajectory if no dark pool orders have been executed before time t_i
$\bar{X}(t_i, t_j)$	upper and lower bounds of the trading regimes for optimal liquidation with adverse selection in discrete time
$A(t_i)$	coefficient matrix for the optimal order in the primary venue

$B(t_i)$	coefficient matrix for the optimal order in the dark pool
$C(t_i)$	value function matrix
$D(t_i) := C(t_i) - C(t_i)\hat{P}\check{C}(t_i)^\dagger \hat{P} C(t_i)$	
$A(t_i, p)$, $B(t_i, p)$, $C(t_i, p)$	respective objects for $n = 1$ and probability of execution p
$A_1(t_i, t_j)$, $A_2(t_i, t_j)$	coefficients of the optimal order in the primary venue for optimal liquidation with adverse selection in discrete time
$B_1(t_i, t_j)$, $B_2(t_i, t_j)$	coefficients of the optimal dark pool order for optimal liquidation with adverse selection in discrete time
$C_1(t_i, t_j)$, $C_2(t_i, t_j)$, $C_3(t_i, t_j)$	value function coefficients for optimal liquidation with adverse selection in discrete time
$\xi(t)$	trading intensity in the primary venue in continuous time
$\eta(t)$	dark pool order in continuous time
$u = (\xi, \eta)$	trading strategy in continuous time
l	end-costs parameter
l_0, l_1	lower bounds for the end-costs parameter
$u^*(l) = (\xi^*(l), \eta^*(l))$	optimal trading strategy in continuous time for finite end-costs
$u^* = (\xi^*, \eta^*)$	optimal liquidation strategy in continuous time
$X^u(t)$	portfolio position associated with the trading strategy u
$X^*(l, t)$	portfolio position associated with the trading strategy $u^*(l)$
$X^*(t)$	portfolio position associated with the trading strategy u^*
$\bar{X}(t, s)$	optimal trading trajectory associated with a certain initial asset position
$g(t, x)$	interpolation function
$\tilde{I}$	diagonal matrix with 1's and 0's on the diagonal depending on whether $\theta_i > 0$
$C(l, t)$	value function matrix for finite end-costs
$C(t)$	value function matrix for optimal liquidation
$\tilde{C} := \operatorname{diag}\left(\frac{\theta_i}{c_{i,i}}\right)$	
$\bar{C} := \operatorname{diag}\left(\frac{1}{c_{i,i}}\right)$	
$C(t; \theta)$	value function coefficient for $n = 1$ and intensity θ
$C_1(t, s)$, $C_2(t, s)$, $C_3(t, s)$	value function coefficients for optimal liquidation with adverse selection in continuous time

$\mu(s) := \frac{2C(s;0)+\theta\Lambda}{\tilde{\theta}\Lambda}$	
$\kappa(s) := \operatorname{arcoth}(\mu(s))$	
$s \wedge t := \min\{s, t\}$	
HJB equation	Hamilton-Jacobi-Bellman equation

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 11.02.2011 Peter Kratz